# Normalization in Lie algebras via mould calculus and applications

Thierry Paul, David Sauzin

# NORMALIZATION IN LIE ALGEBRAS VIA MOULD CALCULUS AND APPLICATIONS

THIERRY PAUL AND DAVID SAUZIN

Abstract. We establish Écalle's mould calculus in an abstract Lie-theoretic setting and use it to solve a normalization problem, which covers several formal normal form problems in the theory of dynamical systems. The mould formalism allows us to reduce the Lie-theoretic problem to a mould equation, the solutions of which are remarkably explicit and can be fully described by means of a gauge transformation group.

The dynamical applications include the construction of Poincaré-Dulac formal normal forms for a vector field around an equilibrium point, a formal infinite-order multiphase averaging procedure for vector fields with fast angular variables (Hamiltonian or not), or the construction of Birkhoff normal forms both in classical and quantum situations. As a by-product we obtain, in the case of harmonic oscillators, the convergence of the quantum Birkhoff form to the classical one, without any Diophantine hypothesis on the frequencies of the unperturbed Hamiltonians.

Contents



*Date*: April 22, 2016.

## Introduction

We are interested in the following situation: given $X_0, B \in \mathcal{L}$, where $\mathcal{L}$ is a Lie algebra over a field $\mathbf{k}$ of characteristic zero, we look for a Lie algebra automorphism $\Psi$ which maps $X_0 + B$ to an element of $\mathcal{L}$ which commutes with $X_0$. We call such a $\Psi$ a "normalizing automorphism" and $\Psi(X_0 + B)$ is then called a "normal form" of $X_0 + B$. Our key assumption will be that $B$ can be decomposed into a sum $B = \sum B_n$ of eigenvectors of the inner derivation $\operatorname{ad}_{X_0}\colon Y \mapsto [X_0, Y]$. We will also assume that $\mathcal{L}$ is a "complete filtered Lie algebra" (Definition 1.1 below), which will allow us to look for $\Psi$ in the form of the exponential of an auxiliary inner derivation.

Our first aim in this article is to introduce Écalle's "mould calculus" for this situation, in the simplest possible way, and to use it to find an explicit solution to the normalization problem: we will obtain $\Psi = \exp(\operatorname{ad}_Y)$ and $\Psi(X_0 + B) = X_0 + Z$ with $Y, Z \in \mathcal{L}$ given by explicit formal series involving all possible iterated Lie brackets $[B_{n_r}, [\dots[B_{n_2}, B_{n_1}]\dots]]$. It is the family of coefficients that one puts in front of these iterated Lie brackets that is called a "mould"; we shall be led to an equation for the moulds associated with $Y$ and $Z$, and our second main result will consist in describing all its solutions, especially all those which are "alternal moulds" (see below), and giving an algorithm to compute them.

Next, we give applications of our result to perturbation theory in classical and quantum dynamics. Indeed, there are several formal normalization problems for dynamical systems or quantum systems which can be put in the above form:

- the construction of Poincaré-Dulac formal normal forms for a vector field around an equilibrium point with diagonalizable linear part, taking for $X_0$ the linear part of the vector field and for $\mathcal{L}$ the Lie algebra of formal vector fields;
- the construction of Hamiltonian Birkhoff normal forms at an elliptic equilibrium point, taking for $X_0$ the quadratic part of the Hamiltonian and for $\mathcal{L}$ the Poisson algebra of formal Hamiltonian functions;
- the elimination at every perturbative order ("averaging") of a fast angular variable $\varphi \in \mathbb{T}^d$ with fixed frequency $\omega \in \mathbb{R}^d$ in a slow-fast vector field (Hamiltonian or not), taking $X_0 = \sum \omega_j \frac{\partial}{\partial \varphi_j}$;
- the construction of quantum Birkhoff normal forms in a Rayleigh-Schrödinger-type situation, taking for $X_0$ the unperturbed part of the quantum Hamiltonian and for $\mathcal{L}$ a Lie algebra of operators of the underlying Hilbert space.

| | Nature of the Lie algebra $\mathcal{L}$ and its Lie bracket | Element to be normalized $X = X_0 + B$, $B = \sum B_n$ | Normalization $e^{\mathrm{ad}_Y} X = X_0 + Z$ |
|---|---|---|---|
| Poincaré-Dulac normal form | Formal vector fields in $z_1, \dots, z_N$ with their natural Lie bracket | $X_0 = \sum_{j=1}^{N} \omega_j z_j \partial_{z_j}$ <br> $B = \sum_{n \in \mathcal{N}} B_n$ <br> $\mathcal{N} = \{ \langle k, \omega \rangle - \omega_j \} \subset \mathbb{C}$ | $e^{\mathrm{ad}_Y} X = \Phi_*^{-1} X$, <br> $\Phi :=$ formal time-1 map for $Y$, <br> $Z$ resonant |
| Birkhoff normal form | Formal Hamiltonians in $x_1, y_1, \dots, x_d, y_d$ with Poisson bracket for $\sum \mathrm{d}x_j \wedge \mathrm{d}y_j$ | $X_0 = \sum_{j=1}^{d} \frac{1}{2} \omega_j (x_j^2 + y_j^2)$ <br> $B = \sum_{n \in \mathbb{Z}^d} B_n$ <br> $\lambda(n) = \mathrm{i} \langle n, \omega \rangle$ | $e^{\mathrm{ad}_Y} X = X \circ \Phi$, <br> $\Phi :=$ formal time-1 map for the Hamiltonian vector field $\{Y, \cdot\}$, <br> $Z$ resonant |
| Multiphase averaging | Vector fields or Hamiltonians $\sum F_j \frac{\partial}{\partial \varphi_j} + \sum G_k \frac{\partial}{\partial I_k}$ or $H(\varphi, I)$ trigonometric polyn. in $\varphi$, smooth in $I$, formal in $\varepsilon$ | $X_0 = \sum \omega_j \frac{\partial}{\partial \varphi_j}$ or $\langle \omega, I \rangle$ <br> $B = \sum_{n \in \mathbb{Z}^d} B_n$ <br> $\lambda(n) = \mathrm{i} \langle n, \omega \rangle$ | $e^{\mathrm{ad}_Y} X = \Phi_*^{-1} X$ or $X \circ \Phi$, <br> $\Phi :=$ formal time-1 map for $Y$ or $\{Y, \cdot\}$, <br> $Z$ resonant, formal in $\varepsilon$ |
| Quantum perturbation theory | $\mathcal{L}^{\mathbb{C}}_{\mathbf{e}}[[\varepsilon]]$, operators in a Hilbert formal in $\varepsilon$, finite-column w.r.t. an orthonormal basis $\mathbf{e}$, $[\cdot\,,\cdot]_{\mathrm{qu}} := \frac{1}{\mathrm{i}\hbar} \times$ commutator | $X_0 = \sum_{k \in I} \lvert e_k \rangle E_k \langle e_k \rvert$ <br> $B = \sum_{n \in \mathcal{N}} B_n$ <br> $\mathcal{N} = \{ \frac{1}{\mathrm{i}\hbar}(E_\ell - E_k) \} \subset \mathbb{C}$ | $e^{\mathrm{ad}_Y} X = e^{\frac{1}{\mathrm{i}\hbar} Y} X e^{-\frac{1}{\mathrm{i}\hbar} Y}$, <br> $Z$ block-diagonal on $\mathbf{e}$ and formal in $\varepsilon$ |
| Quantum perturbation theory uniform in $\hbar \to 0$ | $\mathcal{L}^{\mathbb{R}}_{\mathbf{e},\mathrm{fb}}[[\varepsilon]]$, operators in $L^2(\mathbb{R}^d)$ obtained by Weyl quantization $[\cdot\,,\cdot]_{\mathrm{qu}} := \frac{1}{\mathrm{i}\hbar} \times$ commutator | $X_0 =$ $-\frac{1}{2}\hbar^2 \Delta_{\mathbb{R}^d} + \sum \frac{1}{2} \omega_j^2 x_j^2$ <br> $B = \sum_{n \in \mathbb{Z}^d} B_n$ <br> $\lambda(n) = \mathrm{i} \langle n, \omega \rangle$ | $e^{\mathrm{ad}_Y} X = e^{\frac{1}{\mathrm{i}\hbar} Y} X e^{-\frac{1}{\mathrm{i}\hbar} Y}$, <br> symbol of $Z$ tending to classical B.N.F. as $\hbar \to 0$ |

Table 1. Synthetic overview of applications to dynamics

There is a fifth application, dealing with the way the coefficients of the quantum Birkhoff normal forms formally converge, as $\hbar \to 0$, to those of the classical Birkhoff normal form.

The reader will find a synthetic overview of the dynamical applications in Table 1 on p. 3 and more explanations in Sections 5–9, particularly about the way one can use "homogeneity" to decompose a given $B$ into a sum $\sum B_n$ of eigenvectors of $\mathrm{ad}_{X_0}$ (the indices $n$ belong to a countable set depending on the chosen example; the eigenvalue associated with $n$ is denoted by $\lambda(n)$ when it is not $n$ itself).

In our view, one of the merits of the Lie-theoretic framework we have devised is its unifying power. Indeed, the dynamical applications we have mentioned are well-known, but what is new is the way we obtain each of them as a by-product of one theorem on the normalization problem in a Lie algebra which itself derives from one theorem on the solutions of a certain mould equation. The fact that one can use exactly the same moulds in all these applications is in itself remarkable. This point of view offers a better understanding of the combinatorics involved in these applications. In particular we shall see that our approach gives a more direct way of relating quantum and classical normal forms (last line of Table 1).

Normal forms in completed graded Lie algebras have been studied in [Men13], which is dedicated to logarithmic derivatives associated with graded derivations, motivated by perturbative quantum field theory. However, we see no obvious way of deducing our main results from [Men13], which works in a different context and adopts a more Hopf-algebraic point of view without involving any moulds.

A forthcoming paper [PS16] will be devoted to normal form problems similar to the ones studied in the present article (including applications to classical and quantum dynamics), but in the framework of *Banach scales of Lie algebras*; there, the focus will be on more quantitative results, which can be obtained thanks to the mould representation of the solution in a more analytic context.

Our method relies on Écalle's concept of "alternal mould" ([Eca81], [Eca93]) and owes a lot to the article [EV95] (particularly the part on the so-called "mould of the regal prenormal form" in it). Our approach is however slightly different, and it incorporates a more direct introduction of alternality, because we work in a Lie algebra rather than with an associative algebra of operators which would themselves act on an associative algebra. We do not require from the reader any previous knowledge of the mould formalism. We will provide original self-contained proofs, except for a few elementary facts of Écalle's theory the proof of which can be found e.g. in [Sau09]; at a technical level, we shall use crucially the "dimoulds" introduced in [Sau09].

The core of our work consists in finding and describing the alternal moulds solutions to a certain equation. This is tightly related to algebraic combinatorics. For instance, finite-support alternal moulds can be identified with the primitive elements of a certain combinatorial Hopf algebra, and general alternal moulds with the infinitesimal characters of the dual Hopf algebra. Moreover, the mould counterpart to the grouplike elements of this Hopf algebra and the characters of its dual is embodied in Écalle's concept of "symmetrality". Solving our mould equation will lead us to a generalisation of the classical character of the combinatorial Hopf algebra QSym related to the Dynkin Lie idempotent. However, in this article, we shall not use the language of Hopf algebras but rather stick to Écalle's mould calculus and its application to our Lie-theoretic problem.

The article is divided into three parts.

- The part "Main general results" contains two sections. The first is devoted to the statement of the first main result, Theorem A, in the context of complete filtered Lie algebras. The second section gives the minimum amount of the mould formalism necessary to state the second main result, Theorem B, about the set of all alternal solutions to a certain mould equation.
- The part "Lie mould calculus" contains two sections: Section 3 explains the origin of the notion of alternal mould in relation with computations in a Lie algebra, and then derives the proof of Theorem A from Theorem B. Section 4 gives the proof of Theorem B with the help of "dimoulds".
- The part "Five dynamical applications" contains five sections, each devoted to a particular application of Theorem A: Section 5 for Poincaré-Dulac normal forms of formal vector fields, Section 6 for classical Birkhoff normal forms of formal Hamiltonians, Section 7 for the elimination of a fast angular phase in formal slow-fast vector fields, Section 8 for quantum Birkhoff normal forms of formal perturbations of certain quantum Hamiltonians, Section 9 for the formal convergence of quantum Birkhoff normal forms to classical Birkhoff normal forms as $\hbar \to 0$ for perturbations of harmonic oscillators. To our knowledge, the latter result, valid for arbitrary frequencies, is new and generalizes earlier ones [GP87] [DGH91], which required a Diophantine condition. These applications, though more specialized than the main general results, are written in a self-contained way so as to be (hopefully) accessible to readers who are not specialists of the different domains they cover.

# Main general results

## 1. **Normalization in complete filtered Lie algebras (Theorem A)**

Throughout the article we use the notations

$$\mathbb{N} = \{0, 1, 2, \ldots\}, \qquad \mathrm{i} = \sqrt{-1}.$$

**Definition 1.1.** A "complete filtered Lie algebra" is a Lie algebra $\big(\mathcal{L}, [\,.\,,.\,]\big)$ together with a sequence of subspaces

$$\mathcal{L} = \mathcal{L}_{\geqslant 0} \supset \mathcal{L}_{\geqslant 1} \supset \mathcal{L}_{\geqslant 2} \supset \ldots \qquad \text{with } [\mathcal{L}_{\geqslant m}, \mathcal{L}_{\geqslant n}] \subset \mathcal{L}_{\geqslant m+n} \text{ for all } m, n \in \mathbb{N}$$

(exhaustive decreasing filtration compatible with the Lie bracket) such that $\bigcap \mathcal{L}_{\geqslant m} = \{0\}$ (the filtration is separated) and $\mathcal{L}$ is a complete metric space for the distance $d(X, Y) := 2^{-\operatorname{ord}(Y-X)}$, where we denote by $\operatorname{ord}\colon \mathcal{L} \to \mathbb{N} \cup \{\infty\}$ the order function associated with the filtration (function characterized by $\operatorname{ord}(X) \geqslant m \Leftrightarrow X \in \mathcal{L}_{\geqslant m}$).

The completeness assumption will be used as follows: given a set $I$, a family $(Y_i)_{i \in I}$ of $\mathcal{L}$ is said to be "formally summable" if, for any $m \in \mathbb{N}$, the set $\{\, i \in I \mid Y_i \notin \mathcal{L}_{\geqslant m} \,\}$ is finite; one can then check that the support of this family is countable (if not $I$ itself) and that, for any exhaustion $(I_k)_{k \in \mathbb{N}}$ of this support by finite sets, the sequence $\sum_{i \in I_k} Y_i$ is Cauchy, with a limit which is independent of the exhaustion—this common limit is simply denoted by $\sum_{i \in I} Y_i$.

Here is a simple and useful example of a formally summable series of operators in $\mathcal{L}$: for any $Y \in \mathcal{L}_{\geqslant 1}$ and $r \in \mathbb{N}$, the operator $(\operatorname{ad}_Y)^r$ maps $\mathcal{L}$ in $\mathcal{L}_{\geqslant r}$, hence, for every $X \in \mathcal{L}$, the series $\mathrm{e}^{\operatorname{ad}_Y}(X) := \sum_{r=0}^{\infty} \frac{1}{r!} (\operatorname{ad}_Y)^r (X)$ is formally summable in $\mathcal{L}$. This allows us to define the operator $\mathrm{e}^{\operatorname{ad}_Y}$, which is a Lie algebra automorphism because $\operatorname{ad}_Y$ is a Lie algebra derivation.

Our first main result is

**Theorem A.** *Let $\mathbf{k}$ be a field of characteristic zero. There exist families of coefficients*

$$F^{\lambda_1, \ldots, \lambda_r},\ G^{\lambda_1, \ldots, \lambda_r} \in \mathbf{k} \quad \textit{for } r \geqslant 1,\ \lambda_1, \ldots, \lambda_r \in \mathbf{k}, \tag{1.1}$$

*explicitly computable by induction on $r$, which satisfy the following: given a complete filtered Lie algebra $\mathcal{L}$ over $\mathbf{k}$ and $X_0 \in \mathcal{L}$, given a set $\mathcal{N}$ and a formally summable family $(B_n)_{n \in \mathcal{N}}$ of $\mathcal{L}$ such that each $B_n$ has order $\geqslant 1$ and is an eigenvector of $\operatorname{ad}_{X_0}$, one has*

$$[X_0, Z] = 0, \qquad \mathrm{e}^{\operatorname{ad}_Y}\Big(X_0 + \sum_{n \in \mathcal{N}} B_n\Big) = X_0 + Z, \tag{1.2}$$

*where $Z, Y \in \mathcal{L}_{\geqslant 1}$ are defined as the following sums of formally summable families:*

$$Z = \sum_{r \geqslant 1} \sum_{n_1, n_2, \ldots, n_r \in \mathcal{N}} \frac{1}{r} F^{\lambda(n_1), \lambda(n_2), \ldots, \lambda(n_r)} [B_{n_r}, [\ldots [B_{n_2}, B_{n_1}] \ldots]] \tag{1.3}$$

$$Y = \sum_{r \geqslant 1} \sum_{n_1, n_2, \ldots, n_r \in \mathcal{N}} \frac{1}{r} G^{\lambda(n_1), \lambda(n_2), \ldots, \lambda(n_r)} [B_{n_r}, [\ldots [B_{n_2}, B_{n_1}] \ldots]] \tag{1.4}$$

*with*

$$\lambda \colon \mathcal{N} \to \mathbf{k}, \qquad \lambda(n) := \textit{eigenvalue of } B_n. \tag{1.5}$$

The proof of Theorem A is in Section 3.4.

As we shall see, the families $F^\bullet = (F^{\lambda_1, \ldots, \lambda_r})$ and $G^\bullet = (G^{\lambda_1, \ldots, \lambda_r})$ are not unique, but $(F^\bullet, G^\bullet)$ is in one-to-one correspondence with an auxiliary family called *gauge generator*, which can be chosen arbitrarily among resonant alternal moulds (see the definitions in Section 2). We will see that, for any choice of the gauge generator, one has $F^{\lambda_1, \ldots, \lambda_r} = 0$ whenever $\lambda_r + \cdots + \lambda_1 \neq 0$ and

$$\lambda_r(\lambda_r + \lambda_{r-1}) \cdots (\lambda_r + \cdots + \lambda_2) \neq 0, \ \ \lambda_r + \cdots + \lambda_1 = 0 \quad \Rightarrow \quad F^{\lambda_1, \ldots, \lambda_r} = \frac{1}{\lambda_r(\lambda_r + \lambda_{r-1}) \cdots (\lambda_r + \cdots + \lambda_2)}. \tag{1.6}$$

The formulas are much more complicated when the denominator vanishes, but there still is an explicit algorithm to compute every coefficient $F^{\lambda_1, \ldots, \lambda_r}$ or $G^{\lambda_1, \ldots, \lambda_r}$ depending on the chosen gauge generator: see formulas (2.14)–(2.17) in Section 2.

**Remark 1.2.** We may accept 0 as an eigenvector, i.e. some of the $B_n$'s may vanish and $\lambda(n)$ need not be specified for those values of $n$. Since the support of a summable family is at most countable, one can always choose

$$\mathcal{N} = \mathbb{N}^* \tag{1.7}$$

without loss of generality (by numbering the support of $(B_n)$ and, if this support is finite, setting $B_n = 0$ for the extra values of $n$). On the other hand, one can decide to group together the eigenvectors associated with the same eigenvalue and take for $\mathcal{N}$ the countable subset of $k$ consisting of the eigenvalues which appear in the problem, in which case

$$\mathcal{N} \subset \mathbf{k}, \qquad \lambda(n) = n \ \text{ for } n \in \mathcal{N} \tag{1.8}$$

(this latter choice is the one of [EV95]). In this article we do not opt for any of these two choices and simply consider a general eigenvalue map (1.5) with arbitrary $\mathcal{N}$ (without assuming $B_n \neq 0$ for each $n$).

**Remark 1.3.** The factor $\frac{1}{r}$ in (1.3)–(1.4) is just a convenient normalization. We shall see in Section 3.5 that the inner derivation $\operatorname{ad}_Y$ itself can be written

$$\operatorname{ad}_Y = \sum_{r\geqslant 1} \sum_{n_1,n_2,\dots,n_r\in\mathcal{N}} \frac{1}{r} G^{\lambda(n_1),\lambda(n_2),\dots,\lambda(n_r)} [\operatorname{ad}_{B_{n_r}}, [\dots [\operatorname{ad}_{B_{n_2}}, \operatorname{ad}_{B_{n_1}}]\dots]]$$

$$= \sum_{r\geqslant 1} \sum_{n_1,\dots,n_r\in\mathcal{N}} G^{\lambda(n_1),\cdots,\lambda(n_r)} \operatorname{ad}_{B_{n_r}} \cdots \operatorname{ad}_{B_{n_1}} \quad (1.9)$$

(no more factor $\frac{1}{r}$ in the last series!—note that in general the individual composite operators $\operatorname{ad}_{B_{n_r}} \cdots \operatorname{ad}_{B_{n_1}}$ are not derivations of $\mathcal{L}$). We shall also define a family of coefficients $S^\bullet$ tightly related to $G^\bullet$ such that

$$e^{\operatorname{ad}_Y} = \operatorname{Id} + \sum_{r\geqslant 1} \sum_{n_1,\dots,n_r\in\mathcal{N}} S^{\lambda(n_1),\cdots,\lambda(n_r)} \operatorname{ad}_{B_{n_r}} \cdots \operatorname{ad}_{B_{n_1}} . \quad (1.10)$$

**Remark 1.4.** If $Z, Y \in \mathcal{L}_{\geqslant 1}$ solve equation (1.2), then any $W \in \mathcal{L}_{\geqslant 1}$ such that $[X_0, W] = 0$ gives rise to a solution $(\tilde{Z}, \tilde{Y})$ by setting $\tilde{Z} := e^{\operatorname{ad}_W} Z$ and $\tilde{Y} := \operatorname{BCH}(W, Y) = W + Y + \frac{1}{2}[W, Y] + \cdots$, the Baker-Campbell-Hausdorff series, which is formally summable and satisfies $e^{\operatorname{ad}_{\tilde{Y}}} = e^{\operatorname{ad}_W} e^{\operatorname{ad}_Y}$.

In Section 6, we shall see an example in which $Z$ is unique but $Y$ is not.

We conclude this section with a "truncated version" of Theorem A:

**Addendum to Theorem A.** *Take $\mathcal{L}$, $X_0$, $(B_n)_{n\in\mathcal{N}}$ and $\lambda\colon \mathcal{N} \to \mathbf{k}$ as in the assumptions of Theorem A. Then, for each $m \in \mathbb{N}^*$, the set $\mathcal{N}_m := \{ n \in \mathcal{N} \mid B_n \notin \mathcal{L}_{\geqslant m}\}$ is finite and the finite sums*

$$Z_m := \sum_{r=1}^{m-1} \sum_{n_1,\dots,n_r\in\mathcal{N}_m} \frac{1}{r} F^{\lambda(n_1),\dots,\lambda(n_r)} [B_{n_r}, [\dots [B_{n_2}, B_{n_1}]\dots]], \quad (1.11)$$

$$Y_m := \sum_{r=1}^{m-1} \sum_{n_1,\dots,n_r\in\mathcal{N}_m} \frac{1}{r} G^{\lambda(n_1),\dots,\lambda(n_r)} [B_{n_r}, [\dots [B_{n_2}, B_{n_1}]\dots]] \quad (1.12)$$

*define $Z_m, Y_m \in \mathcal{L}_{\geqslant 1}$ satisfying $[X_0, Z_m] = 0$ and*

$$e^{\operatorname{ad}_{Y_m}} \Big( X_0 + \sum_{n\in\mathcal{N}} B_n \Big) = X_0 + Z_m \mod \mathcal{L}_{\geqslant m}. \quad (1.13)$$

The proof is in Section 3.6.

## 2. The mould equation and its solutions (Theorem B)

We now describe the part of Écalle's mould formalism which will allow us to construct the aforementioned families of coefficients. This will lead us to an equation, of which we will describe all solutions.

**2.1** Let $\mathbf{k}$ a field and $\mathcal{N}$ a nonempty set, considered as an alphabet. We denote by $\underline{\mathcal{N}}$ the corresponding free monoid, whose elements are called *words*,

$$\underline{\mathcal{N}} := \{\, \underline{n} = n_1 \cdots n_r \mid r \in \mathbb{N},\ n_1, \ldots, n_r \in \mathcal{N} \,\}.$$

The monoid law is word *concatenation*: $\underline{a}\,\underline{b} = a_1 \cdots a_r b_1 \cdots b_s$ for $\underline{a} = a_1 \cdots a_r$ and $\underline{b} = b_1 \cdots b_s$. Its unit is the empty word, denoted by $\varnothing$, the only word of length 0. The length of a word $\underline{n}$ is denoted by $r(\underline{n})$. (Given $r \in \mathbb{N}$, we sometimes identify the set of all words of length $r$ with $\mathcal{N}^r$.)

We call *mould* any map $\underline{\mathcal{N}} \to \mathbf{k}$. It is customary to denote the value of the mould on a word $\underline{n}$ by affixing $\underline{n}$ as an upper index to the symbol representing the mould, and to refer to the mould itself by using a big dot as upper index; hence $M^\bullet$ is the mould, the value of which at $\underline{n}$ is denoted by $M^{\underline{n}}$.

For example, the families of coefficients $F^\bullet, G^\bullet$ referred to in Theorem A can be considered as moulds, taking $\mathcal{N} = \mathbf{k}$ as alphabet. For that reason, from now on, we will write $F^{\lambda_1 \cdots \lambda_r}$ and $G^{\lambda_1 \cdots \lambda_r}$ to denote the individual coefficients rather than $F^{\lambda_1, \ldots, \lambda_r}$ or $G^{\lambda_1, \ldots, \lambda_r}$ as in (1.1).

The set $\mathbf{k}^{\underline{\mathcal{N}}}$ of all moulds is clearly a linear space over $\mathbf{k}$. It is also an associative $\mathbf{k}$-algebra (usually not commutative): *mould multiplication* is induced by word concatenation,

$$P^\bullet = M^\bullet \times N^\bullet \ \text{ is defined by } \ \underline{n} \in \underline{\mathcal{N}} \mapsto P^{\underline{n}} := \sum_{\underline{n} = \underline{a}\,\underline{b}} M^{\underline{a}} N^{\underline{b}} \tag{2.1}$$

(summation over all pairs of words $(\underline{a}, \underline{b})$ such that $\underline{n} = \underline{a}\,\underline{b}$, including $(\underline{n}, \varnothing)$ and $(\varnothing, \underline{n})$, thus there are $r(\underline{n}) + 1$ terms in the sum).[1] The multiplication unit is the elementary mould $1^\bullet$ defined by $1^\varnothing = 1$ and $1^{\underline{n}} = 0$ for $\underline{n} \neq \varnothing$. It is easy to see that a mould $M^\bullet$ is invertible if and only if $M^\varnothing \neq 0$; we then denote its multiplicative inverse by ${}^{\mathrm{inv}}M^\bullet$.

The Lie algebra associated with the associative algebra $\mathbf{k}^{\underline{\mathcal{N}}}$ will be denoted $\mathrm{Lie}(\mathbf{k}^{\underline{\mathcal{N}}})$ (same underlying vector space, with bracketing $[M^\bullet, N^\bullet] := M^\bullet \times N^\bullet - N^\bullet \times M^\bullet$).

The order function $\mathrm{ord}\colon \mathbf{k}^{\underline{\mathcal{N}}} \to \mathbb{N} \cup \{\infty\}$ defined by

$$\mathrm{ord}(M^\bullet) \geqslant m \quad \Leftrightarrow \quad M^{\underline{n}} = 0 \text{ whenever } r(\underline{n}) < m \tag{2.2}$$

allows us to view $\mathbf{k}^{\underline{\mathcal{N}}}$ as a complete filtered associative algebra (because the distance $d(M^\bullet, N^\bullet) := 2^{-\mathrm{ord}(N^\bullet - M^\bullet)}$ makes it a complete metric space and $\mathrm{ord}(M^\bullet \times N^\bullet) \geqslant \mathrm{ord}(M^\bullet) + \mathrm{ord}(N^\bullet)$). We

[1] The linear space $\mathbf{k}^{\underline{\mathcal{N}}}$ can be identified with the dual of $\mathbf{k}\underline{\mathcal{N}}$, the $\mathbf{k}$-vector space consisting of all linear combinations of words (formal sums of the form $\sum x_{\underline{n}}\,\underline{n}$, with finitely many nonzero coefficients $x_{\underline{n}} \in \mathbf{k}$): the mould $M^\bullet$ gives rise to the linear form $x \in \mathbf{k}\underline{\mathcal{N}} \mapsto M^\bullet(x) \in \mathbf{k}$ defined by $M^\bullet(\sum x_{\underline{n}}\,\underline{n}) = \sum x_{\underline{n}} M^{\underline{n}}$. The associative algebra structure on $\mathbf{k}^{\underline{\mathcal{N}}}$ is then dual to the coalgebra structure induced on $\mathbf{k}\underline{\mathcal{N}}$ by "word deconcatenation", for which the coproduct is $\Delta(\underline{n}) = \sum_{\underline{n} = \underline{a}\,\underline{b}} \underline{a} \otimes \underline{b}$.

can thus define the mutually inverse exponential and logarithm maps by the following summable series:

$$M^{\varnothing} = 0 \quad \Rightarrow \quad \mathrm{e}^{M^\bullet} := 1^\bullet + \sum_{k\geqslant 1} \tfrac{1}{k!}(M^\bullet)^{\times k}, \quad \log(1^\bullet + M^\bullet) := \sum_{k\geqslant 1} \tfrac{(-1)^{k-1}}{k}(M^\bullet)^{\times k}.$$

**2.2** Écalle's notion of "alternality" is of fundamental importance. Its motivation will be made clear in Section 3.2. The idea is that, since in the situation of Theorem A we will use a mould $M^\bullet$ as a family of coefficients to be multiplied by iterated Lie brackets (as $F^\bullet$ in (1.3) or $G^\bullet$ in (1.4)), it is natural to impose some symmetry (or, rather, antisymmetry) on the coefficients so as to take into account the antisymmetry of the Lie bracket. For instance, the sum over all two-letter words contains expressions like $\frac{1}{2}M^{n_1n_2}[B_{n_2}, B_{n_1}] + \frac{1}{2}M^{n_2n_1}[B_{n_1}, B_{n_2}]$, which coincide with $\frac{1}{2}(M^{n_1n_2} - M^{n_2n_1})[B_{n_2}, B_{n_1}]$, so it is natural to impose

$$M^{n_1n_2} + M^{n_2n_1} = 0 \quad \text{for all } n_1, n_2 \in \mathcal{N}, \tag{2.3}$$

so as to reduce to 1 the number of degrees of freedom associated with the words $n_1n_2$ and $n_2n_1$. Alternality is a generalisation of (2.3) for all lengths $\geqslant 2$.

The definition of alternality is based on word shuffling. Roughly speaking, the shuffling of two words $\underline{a}$ and $\underline{b}$ is the set[2] of all words obtained by interdigitating the letters of $\underline{a}$ and $\underline{b}$ while preserving their internal order in $\underline{a}$ or $\underline{b}$; the number of different ways a word $\underline{n}$ can be obtained out of $\underline{a}$ and $\underline{b}$ is called shuffling coefficient. We make this more precise by using permutations as follows. For $r \in \mathbb{N}$, we let $\mathfrak{S}_r$ (the symmetric group of degree $r$) act to the right on the set $\mathcal{N}^r$ of all words of length $r$ by

$$\underline{n} = n_1 \cdots n_r \mapsto \underline{n}^\tau := n_{\tau(1)} \cdots n_{\tau(r)} \quad \text{for } \tau \in \mathfrak{S}_r \text{ and } \underline{n} \in \mathcal{N}^r. \tag{2.4}$$

For $0 \leqslant \ell \leqslant r$, we set

$$\underline{n}^\tau_{\leqslant \ell} := n_{\tau(1)} \cdots n_{\tau(\ell)}, \qquad \underline{n}^\tau_{>\ell} := n_{\tau(\ell+1)} \cdots n_{\tau(r)}.$$

We also define

$$\mathfrak{S}_r(\ell) := \{\, \tau \in \mathfrak{S}_r \mid \tau(1) < \cdots < \tau(\ell) \text{ and } \tau(\ell+1) < \cdots < \tau(r) \,\},$$

with the conventions $\mathfrak{S}_r(0) = \mathfrak{S}_r(r) = \{\mathrm{id}\}$.

**Definition 2.1.** Given $\underline{n}, \underline{a}, \underline{b} \in \underline{\mathcal{N}}$, the "shuffling coefficient" of $\underline{n}$ in $(\underline{a}, \underline{b})$ is defined to be

$$\mathrm{sh}\big(\tfrac{\underline{a}, \underline{b}}{\underline{n}}\big) := \mathrm{card}\{\, \tau \in \mathfrak{S}_r(\ell) \mid \underline{n}^\tau_{\leqslant \ell} = \underline{a} \text{ and } \underline{n}^\tau_{>\ell} = \underline{b} \,\}, \quad \text{where } \ell := r(\underline{a}). \tag{2.5}$$

[2] or rather the sum—see footnote 3

For instance, if $n, m, p, q$ are four distinct elements of $\mathcal{N}$,

$$\mathrm{sh}\binom{nmp, mq}{nmqpm} = 0, \qquad \mathrm{sh}\binom{nmp, mq}{nmmqp} = 2, \qquad \mathrm{sh}\binom{nmp, mq}{mnqmp} = 1.$$

**Definition 2.2.** A mould $M^\bullet$ is said to be "alternal" if $M^\varnothing = 0$ and

$$\sum_{\underline{n} \in \underline{\mathcal{N}}} \mathrm{sh}\big(\tfrac{\underline{a},\, \underline{b}}{\underline{n}}\big) M^{\underline{n}} = 0 \quad \text{for any two nonempty words } \underline{a}, \underline{b}. \tag{2.6}$$

For instance, (2.6) with $\underline{a} = n_1$ and $\underline{b} = n_2$ yields (2.3) and, with $\underline{a} = n_1$ and $\underline{b} = n_1 n_2$, it yields

$$2M^{n_1 n_1 n_2} + M^{n_1 n_2 n_1} = 0.$$

Notice that any mould whose support is contained in the set of one-letter words is alternal; so is, in particular, the elementary mould $I^\bullet$ defined by

$$I^{\underline{n}} := \mathbb{1}_{\{r(\underline{n})=1\}} \qquad \text{for any word } \underline{n}. \tag{2.7}$$

We denote by $\mathrm{Alt}^\bullet(\mathcal{N})$ the set of alternal moulds, which is clearly a linear subspace of $\mathbf{k}^{\underline{\mathcal{N}}}$; in fact,[3]

$$\mathrm{Alt}^\bullet(\mathcal{N}) \textit{ is a Lie subalgebra of } \mathrm{Lie}(\mathbf{k}^{\underline{\mathcal{N}}})$$

(see e.g. [Sau09, Prop. 5.1]); this will play a role when returning to the situation of Theorem A.

**2.3** Given a function $\varphi\colon \mathcal{N} \to \mathbf{k}$, we denote by the same symbol $\varphi$ its extension to $\underline{\mathcal{N}}$ as a monoid morphism: $\varphi(\varnothing) = 0$ and

$$\underline{n} = n_1 \cdots n_r \in \underline{\mathcal{N}} \mapsto \varphi(\underline{n}) = \varphi(n_1) + \cdots + \varphi(n_r) \in \mathbf{k} \qquad \text{if } r \geqslant 1. \tag{2.8}$$

The formula

$$\nabla_\varphi\colon\ M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}} \mapsto N^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}, \qquad N^{\underline{n}} := \varphi(\underline{n}) M^{\underline{n}} \quad \text{for any word } \underline{n} \tag{2.9}$$

then defines a derivation of the associative algebra $\mathbf{k}^{\underline{\mathcal{N}}}$ (the Leibniz rule for $\nabla_\varphi$ is an obvious consequence of the identity $\varphi(\underline{a}\,\underline{b}) = \varphi(\underline{a}) + \varphi(\underline{b})$). For example, associated with the constant function $\varphi(n) \equiv 1$ is the derivation $\nabla_1$, which spells

$$\nabla_1 M^{\underline{n}} = r(\underline{n}) M^{\underline{n}} \qquad \text{for any } M \in \mathbf{k}^{\underline{\mathcal{N}}} \text{ and } \underline{n} \in \underline{\mathcal{N}}.$$

[3] Word shuffling gives rise to the "shuffling product", defined by $\underline{a} \,\Delta\, \underline{b} := \sum_{\tau \in \mathfrak{S}_r(\ell)} (\underline{a}\,\underline{b})^{\tau^{-1}} = \sum \mathrm{sh}\big(\tfrac{\underline{a},\, \underline{b}}{\underline{n}}\big)\, \underline{n} \in \mathbf{k}\underline{\mathcal{N}}$ for a pair of words such that $r(\underline{a}) = \ell$ and $r(\underline{a}\,\underline{b}) = r$ and extended to $\mathbf{k}\underline{\mathcal{N}} \times \mathbf{k}\underline{\mathcal{N}}$ by bilinearity, which makes the space $\mathbf{k}\underline{\mathcal{N}}$ of footnote 1 a commutative associative algebra. Alternal moulds can then be identified with the infinitesimal characters of the associative algebra $(\mathbf{k}\underline{\mathcal{N}}, \Delta)$, i.e. when viewed as linear forms they are characterized by $M^\bullet(x \,\Delta\, y) = M^\bullet(x) 1^\bullet(y) + 1^\bullet(x) M^\bullet(y)$. In that point of view, $\mathrm{Alt}^\bullet(\mathcal{N})$ is a Lie subalgebra of $\mathrm{Lie}(\mathbf{k}^{\underline{\mathcal{N}}})$ because $\mathbf{k}\underline{\mathcal{N}}$ is a bialgebra (i.e. there is some kind of compatibility between the deconcatenation coproduct and the shuffling product—in fact, $\mathbf{k}\underline{\mathcal{N}}$ is even a Hopf algebra).

In the situation of Theorem A, the derivation $\nabla_\lambda$ associated with the map (1.5) will play a pre-eminent role. We shall need the following

**Definition 2.3.** Given a map $\lambda\colon \mathcal{N} \to \mathbf{k}$, we call "$\lambda$-resonant" any mould $M^\bullet$ such that $\nabla_\lambda M^\bullet = 0$ and use the notation

$$\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N}) := \{\, M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N}) \mid \nabla_\lambda M^\bullet = 0 \,\}.$$

The "$\lambda$-resonant part" of a mould $M^\bullet$ is denoted by $M^\bullet_{\lambda=0}$ and defined by the formula

$$M^{\underline{n}}_{\lambda=0} := \mathbb{1}_{\{\lambda(\underline{n})=0\}}\, M^{\underline{n}} \qquad \text{for any word } \underline{n}.$$

The "gauge generator" of an alternal mould $M^\bullet$ is defined as

$$\mathscr{J}_\lambda(M^\bullet) := \Big[\mathrm{e}^{-M^\bullet} \times \nabla_1\big(\mathrm{e}^{M^\bullet}\big)\Big]_{\lambda=0}.$$

Note that the space $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ of all $\lambda$-resonant alternal moulds is a Lie subalgebra of $\mathrm{Alt}^\bullet(\mathcal{N})$ (being the kernel of a derivation). Clearly, the $\lambda$-resonant part of a mould is $\lambda$-resonant; a mould $M^\bullet$ is $\lambda$-resonant if and only if $M^\bullet = M^\bullet_{\lambda=0}$ or, equivalently, if and only if $M^{\underline{n}} = 0$ whenever $\lambda(\underline{n}) \neq 0$. We shall see later that the gauge generator of an alternal mould is always alternal and, in fact, $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ coincides with the set of all gauge generators of alternal moulds.

It is worth singling out the particular case of an alphabet contained in $\mathbf{k}$:

**Definition 2.4.** If $\mathcal{N} \subset \mathbf{k}$ and $\lambda\colon \mathcal{N} \to \mathbf{k}$ is the inclusion map, then we use the word "resonant" instead of $\lambda$-resonant, and we use the notations $\nabla$, $\mathrm{Alt}^\bullet_0(\mathcal{N})$, $M^\bullet_0$ and $\mathscr{J}(M^\bullet)$ instead of $\nabla_\lambda$, $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, $M^\bullet_{\lambda=0}$, and $\mathscr{J}_\lambda(M^\bullet)$.

**2.4** We are now in a position to state our second main result, describing all the solutions to a certain mould equation, equation (2.10) below. This result, while being of interest in itself, will yield the main step in the proof of Theorem A. Recall that $I^\bullet$ is the alternal mould defined by (2.7).

**Theorem B.** *Let $\mathcal{N}$ be a nonempty set, $\mathbf{k}$ a field of characteristic zero, and $\lambda\colon \mathcal{N} \to \mathbf{k}$ a map.*
*(i) For every $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, there exists a unique pair $(F^\bullet, G^\bullet)$ of alternal moulds such that*

$$\nabla_\lambda F^\bullet = 0, \qquad \nabla_\lambda\big(\mathrm{e}^{G^\bullet}\big) = I^\bullet \times \mathrm{e}^{G^\bullet} - \mathrm{e}^{G^\bullet} \times F^\bullet, \tag{2.10}$$

$$\mathscr{J}_\lambda(G^\bullet) = A^\bullet. \tag{2.11}$$

*(ii) Suppose that $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ is a solution to equation* (2.10)*. Then the formula*

$$J^\bullet \mapsto (\tilde F^\bullet, \tilde G^\bullet) = \Big(\mathrm{e}^{-J^\bullet} \times F^\bullet \times \mathrm{e}^{J^\bullet},\, \log\big(\mathrm{e}^{G^\bullet} \times \mathrm{e}^{J^\bullet}\big)\Big) \tag{2.12}$$

*establishes a one-to-one correspondence between* $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ *and the set of all solutions* $(\tilde F^\bullet, \tilde G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ *of equation* (2.10)*. Moreover,*

$$\mathscr{I}_\lambda(\tilde G^\bullet) = \mathrm{e}^{-J^\bullet} \times \mathscr{I}_\lambda(G^\bullet) \times \mathrm{e}^{J^\bullet} + \mathrm{e}^{-J^\bullet} \times \nabla_1\big(\mathrm{e}^{J^\bullet}\big). \tag{2.13}$$

The proof of Theorem B is given in Section 4. It is constructive in the sense that we will obtain the following simple algorithm to compute the values of $F^\bullet$ and $S^\bullet := \mathrm{e}^{G^\bullet}$ on any word $\underline{n}$ by induction on its length $r(\underline{n})$: introducing an auxiliary alternal mould $N^\bullet$, one must take

$$S^\varnothing = 1, \qquad F^\varnothing = N^\varnothing = 0 \tag{2.14}$$

and, for $r(\underline{n}) \geqslant 1$,

$$\lambda(\underline{n}) \neq 0 \quad \Rightarrow \qquad F^{\underline{n}} = 0, \quad S^{\underline{n}} = \frac{1}{\lambda(\underline{n})}\Big(S^{`\underline{n}} - \sum_{\underline{n}=\underline{a}\,\underline{b}}^{*} S^{\underline{a}} F^{\underline{b}}\Big), \quad N^{\underline{n}} = r(\underline{n})\, S^{\underline{n}} - \sum_{\underline{n}=\underline{a}\,\underline{b}}^{*} S^{\underline{a}} N^{\underline{b}}, \tag{2.15}$$

$$\lambda(\underline{n}) = 0 \quad \Rightarrow \qquad F^{\underline{n}} = S^{`\underline{n}} - \sum_{\underline{n}=\underline{a}\,\underline{b}}^{*} S^{\underline{a}} F^{\underline{b}}, \quad S^{\underline{n}} = \frac{1}{r(\underline{n})}\Big(A^{\underline{n}} + \sum_{\underline{n}=\underline{a}\,\underline{b}}^{*} S^{\underline{a}} N^{\underline{b}}\Big), \quad N^{\underline{n}} = A^{\underline{n}}, \tag{2.16}$$

where we have used the notation $`\underline{n} := n_2 \cdots n_r$ for $\underline{n} = n_1 n_2 \cdots n_r$ and the symbol $\sum^*$ indicates summation over non-trivial decompositions (i.e. $\underline{a}, \underline{b} \neq \varnothing$ in the above sums); we will see that the mould $F^\bullet$ thus inductively defined is alternal and that

$$G^\varnothing = 0, \qquad G^{\underline{n}} = \sum_{k=1}^{r(\underline{n})} \frac{(-1)^{k-1}}{k} \sum_{\underline{n}=\underline{a}^1\cdots\underline{a}^k}^{*} S^{\underline{a}^1} \cdots S^{\underline{a}^k} \quad \text{for } \underline{n} \neq \varnothing \tag{2.17}$$

then defines the alternal mould $G^\bullet$ which solves (2.10)–(2.11).

**2.5** A few remarks are in order.

*2.5.1.* Given alphabets $\mathcal{M}$ and $\mathcal{N}$, any map $\varphi\colon \mathcal{N} \to \mathcal{M}$ induces a map $\varphi^*\colon M^\bullet \in \mathbf{k}^{\underline{\mathcal{M}}} \mapsto M^\bullet_\varphi \in \mathbf{k}^{\underline{\mathcal{N}}}$ defined by $M_\varphi^{n_1\cdots n_r} := M^{\varphi(n_1)\cdots\varphi(n_r)}$, which is a morphism of associative algebras, mapping $\mathrm{Alt}^\bullet(\mathcal{M})$ to $\mathrm{Alt}^\bullet(\mathcal{N})$ and satisfying $\nabla_{\mu\circ\varphi} \circ \varphi^* = \nabla_\mu$ for any $\mu\colon \mathcal{M} \to \mathbf{k}$. Let $\lambda := \mu \circ \varphi$; one can easily check that, if $A^\bullet \in \mathrm{Alt}^\bullet_{\mu=0}(\mathcal{M})$, then the unique solution $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{M}) \times \mathrm{Alt}^\bullet(\mathcal{M})$ of

$$\nabla_\mu F^\bullet = 0, \qquad \nabla_\mu\big(\mathrm{e}^{G^\bullet}\big) = I^\bullet \times \mathrm{e}^{G^\bullet} - \mathrm{e}^{G^\bullet} \times F^\bullet$$

such that $\mathscr{I}_\mu(G^\bullet) = A^\bullet$ is mapped by $\varphi^*$ to the unique solution in $\mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ of (2.10) with gauge generator $\varphi^*(A^\bullet)$.

*2.5.2.* Let us call "canonical case" the case when $\mathcal{N} = \mathbf{k}$ and $\lambda =$ the identity map. We shall see in Section 3.4 that the moulds $F^\bullet, G^\bullet \in \mathbf{k}^{\underline{\mathbf{k}}}$ which are referred to in Theorem A and give rise to solutions $(Z, Y)$ of equation (1.2) are the ones given by Theorem B in the canonical case with arbitrary $A^\bullet \in \mathrm{Alt}_0^\bullet(\mathbf{k})$. The mould $S^\bullet$ referred to in Remark 1.3 is then $\mathrm{e}^{G^\bullet}$.

We shall see that, with the notations of Theorem B(ii), any $J^\bullet \in \mathrm{Alt}_0^\bullet(\mathbf{k})$ gives rise to $W \in \mathcal{L}_{\geqslant 1}$ such that $[X_0, W] = 0$ and the solution $(\tilde{Z}, \tilde{Y})$ of (1.2) associated with $(\tilde{F}^\bullet, \tilde{G}^\bullet)$ is given by $\tilde{Z} = \mathrm{e}^{\mathrm{ad}_W} Z$ and $\tilde{Y} = \mathrm{BCH}(W, Y)$, in line with Remark 1.4.

*2.5.3.* In part (i) of the statement, one may choose $A^\bullet = 0$; this yields for $(F^\bullet, G^\bullet)$ what we call the "zero gauge solution of equation (2.10)". In the canonical case, the zero gauge solution corresponds to what is treated in [EV95] under the name "royal prenormal form". The rest of the statement and the whole proof given in Section 4 are new.

As a consequence of the remark in Section 2.5.1, the zero gauge solution in the general case $\lambda\colon \mathcal{N} \to \mathbf{k}$ is obtained from the zero gauge solution in the canonical case by applying $\lambda^*$.

*2.5.4.* Another possible normalization aimed at singling out a specific solution of (2.10) in $\mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ consists in requiring $G^\bullet_{\lambda=0} = 0$ (instead of requiring $\mathscr{J}_\lambda(G^\bullet) = 0$). There is a unique such solution and here is how one can see it.

According to the Baker-Campbell-Hausdorff formula, for arbitrary $G^\bullet, J^\bullet \in \big(\mathbf{k}^{\underline{\mathcal{N}}}\big)_{\geqslant 1}$ (i.e. such that $G^\varnothing = J^\varnothing = 0$), we can write

$$\log(\mathrm{e}^{G^\bullet} \times \mathrm{e}^{J^\bullet}) = G^\bullet + J^\bullet + \mathcal{F}(G^\bullet, J^\bullet), \qquad \mathcal{F}(G^\bullet, J^\bullet) = \tfrac{1}{2}[G^\bullet, J^\bullet] + \cdots \in \big(\mathbf{k}^{\underline{\mathcal{N}}}\big)_{\geqslant 2},$$

where the functional $\mathcal{F}$ satisfies $\operatorname{ord}\big(\mathcal{F}(G^\bullet, \tilde{J}^\bullet) - \mathcal{F}(G^\bullet, J^\bullet)\big) \geqslant \operatorname{ord}(\tilde{J}^\bullet - J^\bullet) + 1$ for all $\tilde{J}^\bullet \in \big(\mathbf{k}^{\underline{\mathcal{N}}}\big)_{\geqslant 1}$ (which is a contraction property for the distance mentioned right after (2.2)) and preserves alternality. Now, given a solution $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ to equation (2.10), in view of part (ii) of Theorem B, we see that finding a solution $(\tilde{F}^\bullet, \tilde{G}^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ of (2.10) such that $\tilde{G}^\bullet_{\lambda=0} = 0$ is equivalent to finding $J^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ such that

$$J^\bullet = -G^\bullet_{\lambda=0} - \big[\mathcal{F}(G^\bullet, J^\bullet)\big]_{\lambda=0}. \tag{2.18}$$

The fixed point equation (2.18) has a unique solution $J^\bullet$ in $\big(\mathbf{k}^{\underline{\mathcal{N}}}\big)_{\geqslant 1}$ (because of the contraction property), which is clearly $\lambda$-resonant, and also alternal (because $\mathcal{F}$ preserves alternality). The uniqueness of the mould $J^\bullet$ entails that the solution $(\tilde{F}, \tilde{G})$ is unique (it does not depend on the auxiliary solution $(F^\bullet, G^\bullet)$ we started with).

*2.5.5.* "Symmetral" moulds can be defined as the elements of

$$\mathrm{Sym}^\bullet(\mathcal{N}) := \{\mathrm{e}^{M^\bullet} \mid M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})\} \tag{2.19}$$

and $\big(\mathrm{Sym}^\bullet(\mathcal{N}), \times\big)$ is a group, in general non-commutative (see e.g. [Sau09, Prop. 5.1]; see also Remark 3.10 below).

Thus, using the change of unknown $S^\bullet = \mathrm{e}^{G^\bullet}$, it is equivalent to look for a solution $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ of equation (2.10) or for a solution $(F^\bullet, S^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ of the equation

$$\nabla_\lambda F^\bullet = 0, \qquad \nabla_\lambda S^\bullet = I^\bullet \times S^\bullet - S^\bullet \times F^\bullet, \tag{2.20}$$

and the gauge generator will then be

$$\mathscr{I}_\lambda(\log S^\bullet) = \left[{}^{\mathrm{inv}}S^\bullet \times \nabla_1 S^\bullet\right]_{\lambda=0}. \tag{2.21}$$

This mould $S^\bullet = \mathrm{e}^{G^\bullet}$ is the one which appears in the algorithm (2.14)–(2.16); there, $N^\bullet$ is the auxiliary mould $N^\bullet = {}^{\mathrm{inv}}S^\bullet \times \nabla_1 S^\bullet$.

*2.5.6.* For any choice of $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, from (2.15)–(2.16), one easily gets

$$\lambda(n_1 n_2 \cdots n_r)\lambda(n_2 \cdots n_r) \cdots \lambda(n_r) \neq 0 \quad \Rightarrow$$
$$F^{n_i \cdots n_r} = 0 \text{ for } i = 1, \ldots, r \text{ and } S^{n_1 \cdots n_r} = \frac{1}{\lambda(n_1 n_2 \cdots n_r)\lambda(n_2 \cdots n_r) \cdots \lambda(n_r)}, \tag{2.22}$$

whence (1.6) follows by (2.16) and Section 2.5.2.

Note that it may happen that $\lambda(\underline{n}) \neq 0$ for every nonempty word $\underline{n}$, in which case $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N}) = \{0\}$ and there is only one solution $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ to equation (2.10), namely $F^\bullet = 0$ and $G^\bullet =$ logarithm of the mould $S^\bullet$ defined by (2.22).

For instance, this is what happens if $\mathcal{N} = \mathbb{N}^*$ (positive integers), $\mathbf{k} = \mathbb{Q}$ and $\lambda =$ the inclusion map $\mathbb{N}^* \hookrightarrow \mathbb{Q}$. Formula (2.22) then reads

$$S^{n_1 \cdots n_r} = \frac{1}{(n_1 + n_2 + \cdots + n_r)(n_2 + \cdots + n_r) \cdots n_r}.$$

In that case, the Hopf algebra $\mathbf{k}\underline{\mathcal{N}}$ evoked in footnote 3 is the combinatorial Hopf algebra QSym of "quasi-symmetric functions" and this mould $S^\bullet$ is related to the so-called Dynkin Lie idempotent, of which we thus get interesting generalisations by considering arbitrary maps $\lambda\colon \mathbb{N}^* \to \mathbb{Q}$ and the corresponding symmetral moulds $S^\bullet$.

The canonical case defined in Section 2.5.2 is the opposite: $\mathrm{Alt}^\bullet_0(\mathbf{k})$ is huge. Choosing a resonant alternal mould $A^\bullet$ amounts to choosing an arbitrary constant in $\mathbf{k}$ for $A^0$ (only possibly nonzero value in length 1), an arbitrary odd function $\mathbf{k} \to \mathbf{k}$ for $\lambda_1 \mapsto A^{\lambda_1(-\lambda_1)}$ in length 2, etc.

*2.5.7.* The exponential map induces a bijection from $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ to the set $\mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$ consisting of all $\lambda$-resonant symmetral moulds, which is a subgroup of $\mathrm{Sym}^\bullet(\mathcal{N})$.

According to part (ii) of Theorem B, given a solution $(F^\bullet, S^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ of (2.20), we thus have a bijection

$$K^\bullet \mapsto (\tilde{F}^\bullet, \tilde{S}^\bullet) = \left({}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet,\, S^\bullet \times K^\bullet\right) \tag{2.23}$$

between $\mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$ and the set of all solutions $(\tilde{F}^\bullet, \tilde{S}^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ of (2.20). The map

$$(F^\bullet, S^\bullet) \mapsto \big({}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet,\, S^\bullet \times K^\bullet\big)$$

is called the "gauge transformation" associated with $K^\bullet \in \mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$.

The group $\mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$ is called the "gauge group" of equation (2.20); it acts to the right freely and transitively by gauge transformations on the space of solutions $\big\{(F^\bullet, S^\bullet)\big\} \subset \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$. Its effect on gauge generators is given by the formula

$$\mathscr{J}_\lambda(\log \tilde{S}^\bullet) = {}^{\mathrm{inv}}K^\bullet \times \mathscr{J}_\lambda(\log S^\bullet) \times K^\bullet + {}^{\mathrm{inv}}K^\bullet \times \nabla_1 K^\bullet. \tag{2.24}$$

*2.5.8.* The identities

$$\mathrm{e}^{-J^\bullet} \times A^\bullet \times \mathrm{e}^{J^\bullet} = (\mathrm{e}^{-\operatorname{ad}_{J^\bullet}})A^\bullet = \sum_{k\geqslant 0} \tfrac{(-1)^k}{k!}(\operatorname{ad}_{J^\bullet})^k A^\bullet, \qquad \mathrm{e}^{-J^\bullet} \times \nabla_\lambda(\mathrm{e}^{J^\bullet}) = \sum_{k\geqslant 0} \tfrac{(-1)^k}{(k+1)!}(\operatorname{ad}_{J^\bullet})^k\, \nabla_\lambda J^\bullet$$

to be seen in Section 3.3 (Propositions 3.8(ii) and 3.9(ii)) show that, for any alternal mould $M^\bullet$, the $\lambda$-resonant mould $\mathscr{J}_\lambda(M^\bullet)$ is alternal, as claimed in the paragraph following Definition 2.3, and that the right-hand side of (2.13) or (2.24) is indeed alternal and $\lambda$-resonant (by replacing $\nabla_\lambda$ with $\nabla_1$ and observing that $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ is invariant by $\operatorname{ad}_{J^\bullet}$ for $J^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$).

One can easily find the gauge transformation which maps the zero gauge solution on any given solution: if a given solution $(F^\bullet, S^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ has gauge generator $A^\bullet = \mathscr{J}_\lambda(\log S^\bullet)$, then one finds the desired gauge transformation in terms of $A^\bullet$ by solving the equation

$$\nabla_1 K^\bullet = K^\bullet \times A^\bullet$$

inductively on word length with initial condition $K^\varnothing = 1$ (the unique solution $K^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$ is clearly $\lambda$-resonant and it turns out that it is also symmetral).

## 3. **Lie mould calculus and proof of Theorem A**

### 3.1. **General setting.**

Let us give ourselves a field $\mathbf{k}$ and a nonempty set $\mathcal{N}$, so that we can consider the associative $\mathbf{k}$-algebra $\mathbf{k}^{\underline{\mathcal{N}}}$ of Section 2. We suppose that we are also given a Lie algebra $\mathcal{L}$ over $\mathbf{k}$ and a family $(B_n)_{n\in\mathcal{N}}$ of $\mathcal{L}$.

Let us consider an associative algebra $\mathcal{A}$ over $\mathbf{k}$ such that $\mathcal{L}$ is a Lie subalgebra of $\mathrm{Lie}(\mathcal{A})$ (we denote by $\mathrm{Lie}(\mathcal{A})$ the Lie algebra over $\mathbf{k}$ with the same underlying vector space as $\mathcal{A}$ and bracketing $[x,y] := xy - yx$). For instance, by the Poincaré-Birkhoff-Witt theorem, we may take for $\mathcal{A}$ the universal enveloping algebra of $\mathcal{L}$.

**Definition 3.1.** The "associative comould" is the family $B_\bullet = (B_{\underline{n}})_{\underline{n}\in\underline{\mathcal{N}}}$ defined by

$$B_{\underline{n}} := B_{n_r} \cdots B_{n_1} \in \mathcal{A}$$

for any word $\underline{n} = n_1 \cdots n_r$, with the convention $B_\varnothing := 1_{\mathcal{A}}$. The "Lie comould" is the family $B_{[\bullet]} = (B_{[\underline{n}]})_{\underline{n}\in\underline{\mathcal{N}}}$ defined by $B_{[\varnothing]} := 0$ and

$$B_{[\underline{n}]} := \mathrm{ad}_{B_{n_r}} \circ \cdots \circ \mathrm{ad}_{B_{n_2}} B_{n_1} = [B_{n_r}, [\ldots [B_{n_2}, B_{n_1}] \ldots]] \in \mathcal{L}$$

for any nonempty word $\underline{n} = n_1 \cdots n_r$, with the convention $B_{[n_1]} = B_{n_1}$ when $r = 1$.

Beware that in general, contrarily to the Lie comould, the associative comould is not a family of $\mathcal{L}$, but only of $\mathcal{A}$. Écalle's mould calculus ([Eca81], [Eca93], [Sau09]) deals with finite or infinite sums of the form $\sum M^{\underline{n}} B_{\underline{n}}$ in the associative algebra $\mathcal{A}$, with arbitrary moulds $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$. In this article, we use the phrase "Lie mould calculus" when restricting our attention to finite or infinite sums of the form $\sum M^{\underline{n}} B_{\underline{n}}$ with *alternal* moulds $M^\bullet$ because, as will be shown in a moment, such expressions can be rewritten $\sum \frac{1}{r(\underline{n})} M^{\underline{n}} B_{[\underline{n}]}$ and thus belong to the Lie algebra $\mathcal{L}$.

The shuffling coefficients of Definition 2.1 allow us to express the Lie comould $B_{[\bullet]}$ in terms of the associative comould $B_\bullet$:

**Lemma 3.2.** *For any nonempty word $\underline{n} \in \underline{\mathcal{N}}$,*

$$B_{[\underline{n}]} = \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})} r(\underline{a})\, \mathrm{sh}\big(\begin{smallmatrix}\underline{a},\,\underline{b}\\ \underline{n}\end{smallmatrix}\big)\, B_{\widetilde{\underline{b}}\,\underline{a}},$$

*where, for an arbitrary word $\underline{b} = b_1 \cdots b_s$, we denote by $\widetilde{\underline{b}}$ the reversed word: $\widetilde{\underline{b}} = b_s \cdots b_1$.*

*Proof.* Let us show by induction on $r$ that

$$\sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})}\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{\widetilde{\underline{b}}\,\underline{a}} = 0, \qquad \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})} r(\underline{a})\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{\widetilde{\underline{b}}\,\underline{a}} = B_{[\,\underline{n}\,]} \tag{3.1}$$

for any word $\underline{n}$ of length $r \geqslant 1$. We denote the first sum by $\mathrm{LHS}(\underline{n})$ and the second by $\mathrm{LHS}'(\underline{n})$, and observe that, as a consequence of (2.5),

$$\mathrm{LHS}(\underline{n}) = \sum_{\ell=0}^{r}\sum_{\tau\in\mathfrak{S}_r(\ell)} (-1)^{r-\ell} B_{\widetilde{\underline{n}^{\tau}_{>\ell}}\,\underline{n}^{\tau}_{\leqslant\ell}}, \qquad \mathrm{LHS}'(\underline{n}) = \sum_{\ell=0}^{r}\sum_{\tau\in\mathfrak{S}_r(\ell)} (-1)^{r-\ell}\,\ell\, B_{\widetilde{\underline{n}^{\tau}_{>\ell}}\,\underline{n}^{\tau}_{\leqslant\ell}} \tag{3.2}$$

For $r=1$, we find $\mathrm{LHS}(n_1) = B_{n_1} - B_{n_1} = 0$ and $\mathrm{LHS}'(n_1) = 1\cdot B_{n_1} - 0\cdot B_{n_1} = B_{[n_1]}$.

Let us assume that $r \geqslant 2$ and (3.1) holds for any word $\underline{n}$ of length $r-1$. Given an arbitrary word $\underline{m}$ of length $r$, we write it as $\underline{m} = \underline{n}\,c$, where $\underline{n} \in \mathcal{N}^{r-1}$ and $c \in \mathcal{N}$. When using (3.1) to compute $\mathrm{LHS}(\underline{m})$ or $\mathrm{LHS}'(\underline{m})$, we see that the last letter of $\underline{m}$ must either go at the end of $\underline{b}$ or at the end of $\underline{a}$, or, more precisely, using (3.2), we see that $\mathfrak{S}_r(\ell)$ can be written as a disjoint union

$$\mathfrak{S}_r(\ell) = \mathfrak{B} \sqcup \mathfrak{A}, \qquad \mathfrak{B} := \{\, \tau \in \mathfrak{S}_r(\ell) \mid \tau(r) = r \,\}, \quad \mathfrak{A} := \{\, \tau \in \mathfrak{S}_r(\ell) \mid \tau(r) < r \,\}$$

(note that $\tau \in \mathfrak{A} \Rightarrow 1 \leqslant \ell < r$ and $\tau(\ell) = r$), and there are bijections $\tau \in \mathfrak{B} \mapsto \tau' \in \mathfrak{S}_{r-1}(\ell)$ and $\tau \in \mathfrak{A} \mapsto \tau^* \in \mathfrak{S}_{r-1}(\ell-1)$ (note that $\mathfrak{A}$ is empty when $\ell = 0$) so that

$$\underline{m}^{\tau}_{\leqslant\ell} = \underline{n}^{\tau'}_{\leqslant\ell} \text{ and } \underline{m}^{\tau}_{>\ell} = \underline{n}^{\tau'}_{>\ell}\,c \text{ for } \tau\in\mathfrak{B}, \qquad \underline{m}^{\tau}_{\leqslant\ell} = \underline{n}^{\tau^*}_{\leqslant\ell-1}\,c \text{ and } \underline{m}^{\tau}_{>\ell} = \underline{n}^{\tau^*}_{>\ell-1} \text{ for } \tau\in\mathfrak{A}$$

(namely $\tau'(i) = \tau(i)$ for $1 \leqslant i \leqslant r-1$, and $\tau^*(i) = \tau(i)$ for $i \leqslant \ell-1$ while $\tau^*(i) = \tau(i+1)$ for $\ell \leqslant i \leqslant r-1$).[4] Therefore

$$\mathrm{LHS}(\underline{m}) = \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b}\,c)}\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{c\,\widetilde{\underline{b}}\,\underline{a}} + \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})}\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{\widetilde{\underline{b}}\,\underline{a}\,c}$$

and, since $B_{c\widetilde{\underline{b}}\,\underline{a}} = B_{\widetilde{\underline{b}}\,\underline{a}}B_c$ and $B_{\widetilde{\underline{b}}\,\underline{a}\,c} = B_c B_{\widetilde{\underline{b}}\,\underline{a}}$, we get $\mathrm{LHS}(\underline{m}) = -\mathrm{LHS}(\underline{n})B_c + B_c\,\mathrm{LHS}(\underline{n}) = 0$ by the induction hypothesis; on the other hand,

$$\begin{aligned}\mathrm{LHS}'(\underline{m}) &= \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b}\,c)} r(\underline{a})\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{c\,\widetilde{\underline{b}}\,\underline{a}} + \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})} r(\underline{a}\,c)\,\mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,B_{\widetilde{\underline{b}}\,\underline{a}\,c} \\ &= -\mathrm{LHS}'(\underline{n})B_c + B_c\big(\mathrm{LHS}'(\underline{n}) + \mathrm{LHS}(\underline{n})\big) = [B_c, B_{[\,\underline{n}\,]}] = B_{[\underline{n}\,c]} = B_{[\,\underline{m}\,]}.\end{aligned}$$

□

[4] Another way of seeing this is to consider the "unshuffling coproduct" on the vector space $\mathbf{k}\underline{\mathcal{N}}$ of footnote 1: this is the linear map $\Delta\colon \mathbf{k}\underline{\mathcal{N}} \to \mathbf{k}\underline{\mathcal{N}} \otimes \mathbf{k}\underline{\mathcal{N}}$ determined by $\Delta(\underline{n}) = \sum \mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\,\underline{a}\otimes\underline{b}$, and the above property amounts to the inductive definition $\Delta(\varnothing) = 0$ and $\Delta(\underline{n}\,c) = \Delta(\underline{n})(\varnothing\otimes c + c\otimes\varnothing)$, where we make use of the non-commutative associative "concatenation product" on $\mathbf{k}\underline{\mathcal{N}}$ or $\mathbf{k}\underline{\mathcal{N}}\otimes\mathbf{k}\underline{\mathcal{N}}$ (in fact, this gives rise to another Hopf algebra structure on $\mathbf{k}\underline{\mathcal{N}}$).

3.2. **Finite mould expansions.**

Let us denote by $\mathbf{k}^{(\underline{\mathcal{N}})}$ the set of finite-support moulds, which is clearly an associative subalgebra of $\mathbf{k}^{\underline{\mathcal{N}}}$. The finiteness condition allows us to define a map with values in $\mathcal{A}$ by means of the associative comould $B_\bullet$:

$$M^\bullet \in \mathbf{k}^{(\underline{\mathcal{N}})} \mapsto M^\bullet B_\bullet := \sum_{\underline{n} \in \underline{\mathcal{N}}} M^{\underline{n}} B_{\underline{n}} \in \mathcal{A}. \tag{3.3}$$

Since $B_{\underline{a}\,\underline{b}} = B_{\underline{b}} B_{\underline{a}}$ for any two words $\underline{a}, \underline{b}$, it is obvious that the map (3.3) is an associative algebra anti-morphism, i.e.

$$(M^\bullet \times N^\bullet) B_\bullet = (N^\bullet B_\bullet)(M^\bullet B_\bullet) \quad \text{for any } M^\bullet, N^\bullet \in \mathbf{k}^{(\underline{\mathcal{N}})}. \tag{3.4}$$

We can also define a map with values in $\mathcal{L}$ by means of the Lie comould $B_{[\,\bullet\,]}$:

$$M^\bullet \in \mathbf{k}^{(\underline{\mathcal{N}})} \mapsto M^\bullet B_{[\,\bullet\,]} := \sum_{\underline{n} \neq \varnothing} \tfrac{1}{r(\underline{n})} M^{\underline{n}} B_{[\,\underline{n}\,]} \in \mathcal{L}. \tag{3.5}$$

**Lemma 3.3.** *Let $M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})$ and let $\Omega$ be an orbit of the action (2.4) of $\mathfrak{S}_r$ for some $r \in \mathbb{N}^*$. Then*

$$\sum_{\underline{n} \in \Omega} M^{\underline{n}} B_{[\,\underline{n}\,]} = r \sum_{\underline{n} \in \Omega} M^{\underline{n}} B_{\underline{n}}. \tag{3.6}$$

*If $M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N}) \cap \mathbf{k}^{(\underline{\mathcal{N}})}$, then*

$$M^\bullet B_\bullet = M^\bullet B_{[\,\bullet\,]}. \tag{3.7}$$

*Proof.* Lemma 3.2 allows us to rewrite the left-hand side of (3.6) as

$$\mathrm{LHS} = \sum_{(\underline{a},\underline{b}) \in \underline{\mathcal{N}} \times \underline{\mathcal{N}}} (-1)^{r(\underline{b})} r(\underline{a}) \Big( \sum_{\underline{n} \in \Omega} \mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big) M^{\underline{n}} \Big) B_{\underline{\tilde{b}}\,\underline{a}}.$$

In view of (2.5), the sum between parentheses is 0 if $\underline{a}\,\underline{b} \notin \Omega$, whereas, if $\underline{a}\,\underline{b} \in \Omega$, it is

$$\sum_{\underline{n} \in \underline{\mathcal{N}}} \mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big) M^{\underline{n}}.$$

According to Definition 2.2, the latter sum is 0 when both $\underline{a}$ and $\underline{b}$ are nonempty, and it is $M^{\underline{a}}$ when $\underline{b} = \varnothing$, hence we end up with LHS $= \sum_{\underline{a} \in \Omega} r(\underline{a}) M^{\underline{a}} B_{\underline{a}}$, which coincides with the right-hand side of (3.6).

To prove (3.7), by linearity we can assume that there is $r \geqslant 1$ such that the support of $M^\bullet$ is contained in $\mathcal{N}^r$. Then we can partition $\mathcal{N}^r$ into orbits:

$$M^\bullet B_\bullet = \sum_{\Omega \in \mathcal{N}^r/\mathfrak{S}_r} \sum_{\underline{n} \in \Omega} M^{\underline{n}} B_{\underline{n}} = \tfrac{1}{r} \sum_{\Omega \in \mathcal{N}^r/\mathfrak{S}_r} \sum_{\underline{n} \in \Omega} M^{\underline{n}} B_{[\,\underline{n}\,]} = \tfrac{1}{r} \sum_{\underline{n} \in \mathcal{N}^r} M^{\underline{n}} B_{[\,\underline{n}\,]} = \sum_{\underline{n} \neq \varnothing} \tfrac{1}{r(\underline{n})} M^{\underline{n}} B_{[\,\underline{n}\,]}.$$

$\square$

**Remark 3.4.** An identity more precise than (3.6) is mentioned in Écalle's works: given a letter $c$ and an orbit $\Omega$ of the action (2.4) of $\mathfrak{S}_r$ for some $r \in \mathbb{N}^*$, let $r_c(\Omega)$ denote the number of occurrences of the letter $c$ in any word of $\Omega$ and let $\Omega_c := \{\,\underline{n} \in \Omega \mid n_1 = c\,\}$; then, for any alternal mould $M^\bullet$,

$$\sum_{\underline{n}\in\Omega_c} M^{\underline{n}} B_{[\underline{n}]} = r_c(\Omega) \sum_{\underline{n}\in\Omega} M^{\underline{n}} B_{\underline{n}}.$$

This is related to the identity

$$B_{[c\,\underline{n}]} = \sum_{(\underline{a},\underline{b})\in\underline{\mathcal{N}}\times\underline{\mathcal{N}}} (-1)^{r(\underline{b})} \operatorname{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big)\, B_{\widetilde{\underline{b}}\,c\,\underline{a}} \quad \text{for any } c \in \mathcal{N} \text{ and } \underline{n} \in \underline{\mathcal{N}}$$

and to the following consequence of alternality:

$$M^{\underline{a}\,c\,\underline{b}} = (-1)^{r(\underline{a})} \sum_{\underline{n}\in\underline{\mathcal{N}}} \operatorname{sh}\big(\tfrac{\widetilde{\underline{a}},\,\underline{b}}{\underline{n}}\big) M^{c\,\underline{n}} \quad \text{for any } c \in \mathcal{N} \text{ and } \underline{a}, \underline{b} \in \underline{\mathcal{N}}$$

(stated as formula (5.26) in [EV95]).

Recall that, as mentioned in Section 2.2, the set $\operatorname{Alt}^\bullet(\mathcal{N})$ of alternal moulds is a Lie subalgebra of $\operatorname{Lie}(\mathbf{k}^{\underline{\mathcal{N}}})$. Let us denote the set of finite-support alternal moulds by

$$\operatorname{Alt}^\bullet_f(\mathcal{N}) := \operatorname{Alt}^\bullet(\mathcal{N}) \cap \mathbf{k}^{(\underline{\mathcal{N}})}.$$

It is obvious that $\operatorname{Alt}^\bullet_f(\mathcal{N})$ is also a Lie subalgebra. In view of (3.7), there is no need to distinguish between the maps (3.3) and (3.5) when restricting to $\operatorname{Alt}^\bullet_f(\mathcal{N})$.

**Proposition 3.5.** *The map $M^\bullet \mapsto M^\bullet B_{[\bullet]}$ induces a Lie algebra anti-morphism $\operatorname{Alt}^\bullet_f(\mathcal{N}) \to \mathcal{L}$, i.e.*

$$[M^\bullet, N^\bullet] B_{[\bullet]} = [N^\bullet B_{[\bullet]}, M^\bullet B_{[\bullet]}] \quad \textit{for any } M^\bullet, N^\bullet \in \operatorname{Alt}^\bullet_f(\mathcal{N}).$$

*Proof.* Using (3.4) and (3.7), we compute $[M^\bullet, N^\bullet] B_{[\bullet]} = [M^\bullet, N^\bullet] B_\bullet = (M^\bullet \times N^\bullet) B_\bullet - (N^\bullet \times M^\bullet) B_\bullet = (N^\bullet B_\bullet)(M^\bullet B_\bullet) - (M^\bullet B_\bullet)(N^\bullet B_\bullet) = [N^\bullet B_\bullet, M^\bullet B_\bullet] = [N^\bullet B_{[\bullet]}, M^\bullet B_{[\bullet]}]$. ☐

**Proposition 3.6.** *Suppose that there are a function $\lambda\colon \mathcal{N} \to \mathbf{k}$ and an $X_0 \in \mathcal{L}$ such that $[X_0, B_n] = \lambda(n) B_n$ for each letter $n$. Then*

$$[X_0, M^\bullet B_\bullet] = (\nabla_\lambda M^\bullet) B_\bullet \quad \textit{and} \quad [X_0, M^\bullet B_{[\bullet]}] = (\nabla_\lambda M^\bullet) B_{[\bullet]} \quad \textit{for any } M^\bullet \in \mathbf{k}^{(\underline{\mathcal{N}})}. \tag{3.8}$$

*Proof.* One easily checks that

$$[X_0, B_{\underline{n}}] = \lambda(\underline{n}) B_{\underline{n}} \quad \text{and} \quad [X_0, B_{[\underline{n}]}] = \lambda(\underline{n}) B_{[\underline{n}]} \quad \text{for any } \underline{n} \in \underline{\mathcal{N}}$$

by induction on $r(\underline{n})$ (because $[X_0, \cdot\,]$ is a derivation of the associative algebra $\mathcal{A}$, as well as derivation of the Lie algebra $\mathcal{L}$), whence (3.8) follows. ☐

### 3.3. **Mould expansions in complete filtered Lie algebras.**

We now assume that $\mathcal{L}$ is a complete filtered Lie algebra and that $(B_n)_{n\in\mathcal{N}}$ is a formally summable family such that each $B_n$ has order $\geqslant 1$. We do not need any auxiliary associative algebra $\mathcal{A}$ such that $\mathcal{L} \subset \operatorname{Lie}(\mathcal{A})$ in this section, except at the end of Remark 3.10.

**Lemma 3.7.** *For each $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$ the family $(\frac{1}{r(\underline{n})} M^{\underline{n}} B_{[\underline{n}]})_{\underline{n}\neq\varnothing}$ is formally summable, hence there is a well-defined extension of the map (3.5) to the set of all moulds (for which we use the same notation):*

$$M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}} \mapsto M^\bullet B_{[\bullet]} := \sum_{\underline{n}\neq\varnothing} \tfrac{1}{r(\underline{n})} M^{\underline{n}} B_{[\underline{n}]} \in \mathcal{L}. \tag{3.9}$$

*This is a $\mathbf{k}$-linear map, compatible with the filtrations of $\mathbf{k}^{\underline{\mathcal{N}}}$ and $\mathcal{L}$ in the sense that, for each $m \in \mathbb{N}$ and $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$,*

$$\operatorname{ord}(M^\bullet) \geqslant m \quad \Rightarrow \quad \operatorname{ord}(M^\bullet B_{[\bullet]}) \geqslant m \tag{3.10}$$

*(with the notation (2.2) for the order function associated with the filtration of $\mathbf{k}^{\underline{\mathcal{N}}}$).*

*Proof.* By assumption, $\mathcal{N}_m := \{\, n \in \mathcal{N} \mid \operatorname{ord}(B_n) < m \,\}$ is finite for each $m \in \mathbb{N}$ and, in view of Definition 1.1, $\operatorname{ord}(B_{[\underline{n}]}) \geqslant r(\underline{n})$ for each $\underline{n} \in \underline{\mathcal{N}}$. This implies that

$$\{\, \underline{n} \in \underline{\mathcal{N}} \mid \operatorname{ord}(B_{[\underline{n}]}) < m \,\} \subset \{\, \underline{n} \in \underline{\mathcal{N}} \mid r := r(\underline{n}) < m \text{ and } n_1, \dots, n_r \in \mathcal{N}_m \,\},$$

which is finite, hence the formal summability follows. The property (3.10) is obvious. □

Note that, if $M^\varnothing = 0$ (as is the case when $M^\bullet$ is alternal), then $\mathrm{e}^{M^\bullet}$ is a well-defined mould and $Y := M^\bullet B_{[\bullet]}$ has order $\geqslant 1$, hence $\mathrm{e}^{\operatorname{ad}_Y}$ is a well-defined Lie algebra automorphism.

**Proposition 3.8.** *(i) The map (3.9) induces a Lie algebra anti-morphism $\operatorname{Alt}^\bullet(\mathcal{N}) \to \mathcal{L}$, i.e.*

$$[M^\bullet, N^\bullet] B_{[\bullet]} = [N^\bullet B_{[\bullet]}, M^\bullet B_{[\bullet]}] \quad \textit{for any } M^\bullet, N^\bullet \in \operatorname{Alt}^\bullet(\mathcal{N}). \tag{3.11}$$

*(ii) If $M^\bullet, N^\bullet \in \operatorname{Alt}^\bullet(\mathcal{N})$, then the mould $\mathrm{e}^{-M^\bullet} \times N^\bullet \times \mathrm{e}^{M^\bullet}$ can be written*

$$\mathrm{e}^{-M^\bullet} \times N^\bullet \times \mathrm{e}^{M^\bullet} = \big(\mathrm{e}^{-\operatorname{ad}_{M^\bullet}}\big) N^\bullet = \sum_{k\geqslant 0} \tfrac{(-1)^k}{k!} (\operatorname{ad}_{M^\bullet})^k N^\bullet$$

*and is alternal, and $Y := M^\bullet B_{[\bullet]}$ satisfies*

$$\mathrm{e}^{\operatorname{ad}_Y}(N^\bullet B_{[\bullet]}) = \big(\mathrm{e}^{-M^\bullet} \times N^\bullet \times \mathrm{e}^{M^\bullet}\big) B_{[\bullet]}.$$

*Proof.* (i) As mentioned in Section (2), the set of all moulds $\mathbf{k}^{\underline{\mathcal{N}}}$ is a complete metric space for the distance $d(M^\bullet, N^\bullet) := 2^{-\operatorname{ord}(N^\bullet - M^\bullet)}$. The map $M^\bullet \mapsto M^\bullet B_{[\bullet]}$ is continuous (and even 1-Lipschitz) by (3.10), and the set of finite-support alternal moulds $\operatorname{Alt}^\bullet_f(\mathcal{N})$ is dense in $\operatorname{Alt}^\bullet(\mathcal{N})$, so (3.11) follows from Proposition 3.5.

(ii) Because of (i), the adjoint representations of $\mathrm{Alt}^\bullet(\mathcal{N})$ and $\mathcal{L}$ are related by

$$M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N}),\ Y = M^\bullet B_{[\bullet]} \quad \Rightarrow \quad \mathrm{ad}_Y(N^\bullet B_{[\bullet]}) = -(\mathrm{ad}_{M^\bullet} N^\bullet)B_{[\bullet]} \text{ for any } N^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N}), \tag{3.12}$$

therefore $\mathrm{e}^{\mathrm{ad}_Y}(N^\bullet B_{[\bullet]}) = \big(\mathrm{e}^{-\mathrm{ad}_{M^\bullet}}(N^\bullet)\big)B_{[\bullet]}$, where $\mathrm{e}^{-\mathrm{ad}_{M^\bullet}}(N^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N})$ is well-defined because $M^\varnothing = 0$, hence $\mathrm{ad}_{M^\bullet}$ increases order in $\mathcal{L}$ by at least one unit and $\mathrm{e}^{-\mathrm{ad}_{M^\bullet}}$ is a well-defined $\mathbf{k}$-linear operator of $\mathrm{Alt}^\bullet(\mathcal{N})$.

In fact, $\mathrm{Alt}^\bullet(\mathcal{N}) \hookrightarrow \mathrm{Lie}(\mathbf{k}^{\underline{\mathcal{N}}})$ and $\mathrm{e}^{-\mathrm{ad}_{M^\bullet}}$ is also a well-defined $\mathbf{k}$-linear operator of $\mathbf{k}^{\underline{\mathcal{N}}}$; as such, it can be written

$$\mathrm{e}^{-\mathrm{ad}_{M^\bullet}} = \mathrm{e}^{-L_{M^\bullet}+R_{M^\bullet}} = \mathrm{e}^{-L_{M^\bullet}} \circ \mathrm{e}^{R_{M^\bullet}},$$

where $L_{M^\bullet}, R_{M^\bullet} \in \mathrm{End}_{\mathbf{k}}\big(\mathbf{k}^{\underline{\mathcal{N}}}\big)$ are the operators of left-multiplication and right-multiplication by $M^\bullet$, which commute. Obviously, $\mathrm{e}^{-L_{M^\bullet}}$ and $\mathrm{e}^{R_{M^\bullet}}$ are the operators of left-multiplication and right-multiplication by $\mathrm{e}^{-M^\bullet}$ and $\mathrm{e}^{M^\bullet}$, hence $\mathrm{e}^{-\mathrm{ad}_{M^\bullet}}(N^\bullet) = \mathrm{e}^{-M^\bullet} \times N^\bullet \times \mathrm{e}^{M^\bullet}$ (the latter identity is sometimes called Hadamard lemma; we gave these details because later we will need again the operators $L_{M^\bullet}$ and $R_{M^\bullet}$). □

**Proposition 3.9.** *Suppose that there are a function $\lambda\colon \mathcal{N} \to \mathbf{k}$ and an $X_0 \in \mathcal{L}$ such that $[X_0, B_n] = \lambda(n)B_n$ for each letter $n$. If $M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})$, then*

*(i) the mould $\nabla_\lambda M^\bullet$ is alternal and*

$$[X_0, M^\bullet B_{[\bullet]}] = (\nabla_\lambda M^\bullet)B_{[\bullet]}, \tag{3.13}$$

*(ii) the mould $\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet})$ can be written*

$$\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet}) = \sum_{k\geqslant 0} \tfrac{(-1)^k}{(k+1)!}(\mathrm{ad}_{M^\bullet})^k \nabla_\lambda M^\bullet$$

*and is alternal, and $Y := M^\bullet B_{[\bullet]}$ satisfies*

$$\mathrm{e}^{\mathrm{ad}_Y} X_0 = X_0 - \big(\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet})\big)B_{[\bullet]}.$$

*Proof.* (i) The identity (3.13) holds for any $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$, as a consequence of (3.8), by continuity of $M^\bullet \mapsto M^\bullet B_{[\bullet]}$ and density of $\mathbf{k}^{(\underline{\mathcal{N}})}$ in $\mathbf{k}^{\underline{\mathcal{N}}}$. It is obvious that $\nabla_\lambda$ preserves alternality.

(ii) We write $\mathrm{e}^{\mathrm{ad}_Y} X_0 - X_0 = \sum_{k\geqslant 0} \frac{1}{(k+1)!}(\mathrm{ad}_Y)^{k+1} X_0 = -\sum_{k\geqslant 0} \frac{1}{(k+1)!}(\mathrm{ad}_Y)^k [X_0, Y]$ with $[X_0, Y] = (\nabla_\lambda M^\bullet)B_{[\bullet]}$ by (3.13), whence $(\mathrm{ad}_Y)^k[X_0, Y] = (-1)^k\big((\mathrm{ad}_{M^\bullet})^k \nabla_\lambda M^\bullet\big)\big)B_{[\bullet]}$ by (3.12). Therefore

$$\mathrm{e}^{\mathrm{ad}_Y} X_0 - X_0 = -(P\,\nabla_\lambda M^\bullet)B_{[\bullet]} \quad \text{with } P := \sum_{k\geqslant 0} \tfrac{(-1)^k}{(k+1)!}(\mathrm{ad}_{M^\bullet})^k \in \mathrm{End}_{\mathbf{k}}\big(\mathbf{k}^{\underline{\mathcal{N}}}\big). \tag{3.14}$$

Note that $P$ is a well-defined $\mathbf{k}$-linear operator of $\mathbf{k}^{\underline{\mathcal{N}}}$ which preserves $\mathrm{Alt}^\bullet(\mathcal{N})$, because $\mathrm{ad}_{M^\bullet}$ increases order in $\mathbf{k}^{\underline{\mathcal{N}}}$ by at least one unit and preserves $\mathrm{Alt}^\bullet(\mathcal{N})$.

On the other hand, as $\nabla_\lambda$ is a derivation of the associative algebra $\mathbf{k}^{\underline{\mathcal{N}}}$, the Leibniz formula applied to $\mathrm{e}^{M^\bullet} = 1^\bullet + \sum_{k\geqslant 0} \frac{1}{(k+1)!}(M^\bullet)^{\times(k+1)}$ yields

$$\nabla_\lambda(\mathrm{e}^{M^\bullet}) = \sum_{k\geqslant 0} \tfrac{1}{(k+1)!} \sum_{p+q=k} (M^\bullet)^{\times p} \times \nabla_\lambda M^\bullet \times (M^\bullet)^{\times q} = \sum_{k\geqslant 0} \tfrac{1}{(k+1)!} \sum_{p+q=k} L^p_{M^\bullet} R^q_{M^\bullet}(\nabla_\lambda M^\bullet),$$

with the same left- and right-multiplication operators $L_{M^\bullet}$ and $R_{M^\bullet}$ as in the end of the proof of Proposition 3.8. Left-multiplication by $\mathrm{e}^{-M^\bullet}$ coincides with the operator $\mathrm{e}^{-L_{M^\bullet}}$, therefore

$$\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet}) = Q\,\nabla_\lambda M^\bullet \quad \text{with } Q := \mathrm{e}^{-L_{M^\bullet}} \sum_{k\geqslant 0} \tfrac{1}{(k+1)!} \sum_{p+q=k} L^p_{M^\bullet} R^q_{M^\bullet} \in \mathrm{End}_{\mathbf{k}}\big(\mathbf{k}^{\underline{\mathcal{N}}}\big). \tag{3.15}$$

Since $\mathrm{ad}_{M^\bullet} = L_{M^\bullet} - R_{M^\bullet}$, we see that $P = Q$ in $\mathrm{End}_{\mathbf{k}}\big(\mathbf{k}^{\underline{\mathcal{N}}}\big)$, as a consequence of the following identity between (commutative) series of two indeterminates:

$$\sum_{k\geqslant 0} \tfrac{(-1)^k}{(k+1)!}(L-R)^k = \mathrm{e}^{-L} \sum_{k\geqslant 0} \tfrac{1}{(k+1)!} \sum_{p+q=k} L^p R^q \in \mathbb{Q}[[L,R]]$$

(which can be checked, since $\mathbb{Q}[[L,R]]$ has no divisor of zero, by multiplying both sides by $L-R$: the left-hand side yields $-\mathrm{e}^{-L+R}+1$ and the right-hand side yields $\mathrm{e}^{-L}\sum_{k\geqslant 0}\frac{1}{(k+1)!}(L^{k+1}-R^{k+1}) = \mathrm{e}^{-L}(\mathrm{e}^L - \mathrm{e}^R)$).

Since $P = Q$, (3.15) shows that $\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet}) = P\,\nabla_\lambda M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})$ (because $\nabla_\lambda M^\bullet$ is alternal and $P$ preserves $\mathrm{Alt}^\bullet(\mathcal{N})$), and (3.14) yields $\mathrm{e}^{\mathrm{ad}_Y} X_0 - X_0 = -\big(\mathrm{e}^{-M^\bullet} \times \nabla_\lambda(\mathrm{e}^{M^\bullet})\big) B_{[\bullet]}$. $\square$

**Remark 3.10.** The set $\mathrm{Sym}^\bullet(\mathcal{N}) \subset \mathbf{k}^{\underline{\mathcal{N}}}$ of symmetral moulds has been defined in (2.19) as the set of all exponentials of alternal moulds. Here is a characterization more in the spirit of Definition 2.2 (the proof of which can be found e.g. in [Sau09, Prop. 5.1]): *A mould $M^\bullet$ is symmetral if and only if*

$$M^\varnothing = 1 \quad \textit{and} \quad \sum_{\underline{n}\in\underline{\mathcal{N}}} \mathrm{sh}\big(\tfrac{\underline{a},\,\underline{b}}{\underline{n}}\big) M^{\underline{n}} = M^{\underline{a}} M^{\underline{b}} \quad \textit{for any two nonempty words } \underline{a}, \underline{b}. \tag{3.16}$$

When identifying $\mathbf{k}^{\underline{\mathcal{N}}}$ with the dual of $\mathbf{k}\underline{\mathcal{N}}$ as in footnotes 1 and 3, we thus identify the symmetral moulds with the characters of the associative algebra $(\mathbf{k}\underline{\mathcal{N}}, \Delta)$, i.e. when viewed as linear forms of $\mathbf{k}\underline{\mathcal{N}}$ they are characterised by $M^\bullet(x \,\Delta\, y) = M^\bullet(x) M^\bullet(y)$. In that point of view, $\mathrm{Sym}^\bullet(\mathcal{N})$ is a group because $\mathbf{k}\underline{\mathcal{N}}$ is a bialgebra.

In the case when $\mathcal{L} \hookrightarrow \mathrm{Lie}(\mathcal{A})$, where $\mathcal{A}$ is a complete filtered associative algebra such that $\mathcal{L}_{\geqslant m} = \mathcal{L} \cap \mathcal{A}_{\geqslant m}$ for each $m$, the map (3.3) extends to an associative algebra anti-morphism $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}} \mapsto M^\bullet B_\bullet \in \mathcal{A}$, compatible with the filtrations of $\mathbf{k}^{\underline{\mathcal{N}}}$ and $\mathcal{A}$, whose restriction to $\mathrm{Alt}^\bullet(\mathcal{N})$ coincide with that of $M^\bullet \mapsto M^\bullet B_{[\bullet]}$. Then

$$M^\varnothing = 0 \quad \Rightarrow \quad \mathrm{e}^{M^\bullet B_\bullet} = \big(\mathrm{e}^{M^\bullet}\big) B_\bullet.$$

In particular, if $M^\bullet$ is alternal, then $\mathrm{e}^{M^\bullet B_{[\bullet]}} = \big(\mathrm{e}^{M^\bullet}\big) B_\bullet$ with $\mathrm{e}^{M^\bullet}$ symmetral.

3.4. **Theorem B implies Theorem A.**

In this section, we take Theorem B for granted and show how Theorem A follows from Lie mould calculus. We thus assume that we are given $\mathcal{N}$ a nonempty set, $\mathbf{k}$ a field of characteristic zero, $\lambda\colon \mathcal{N} \to \mathbf{k}$ a map, $\mathcal{L}$ a complete filtered Lie algebra over $\mathbf{k}$, an element $X_0 \in \mathcal{L}$, and a formally summable family $(B_n)_{n\in\mathcal{N}}$ such that $\operatorname{ord}(B_n) \geqslant 1$ and $[X_0, B_n] = \lambda(n)B_n$ for each $n \in \mathcal{N}$.

Let us consider any of the many solutions $(F^\bullet, G^\bullet) \in \operatorname{Alt}^\bullet(\mathbf{k}) \times \operatorname{Alt}^\bullet(\mathbf{k})$ of equation (2.10) that Theorem B provides in the canonical case of Section 2.5.2, i.e. with $\nabla_{\mathrm{id}}$ replacing $\nabla_\lambda$. We thus have alternal moulds $F^\bullet, G^\bullet$, explicitly defined by (2.14)–(2.17) with some $A^\bullet \in \operatorname{Alt}^\bullet_0(\mathbf{k})$, which satisfy equation (2.10).

Using the map $\lambda^*\colon \mathbf{k}^{\underline{\mathbf{k}}} \to \mathbf{k}^{\underline{\mathcal{N}}}$ of Section 2.5.1, we define $F^\bullet_\lambda := \lambda^*(F^\bullet)$ and $G^\bullet_\lambda := \lambda^*(G^\bullet)$, which belong to $\operatorname{Alt}^\bullet(\mathcal{N})$ and satisfy equation (2.10) but now with the operator $\nabla_\lambda$ associated with the eigenvalue map $\lambda$.

Let $Z := F^\bullet_\lambda B_{[\bullet]}$, in accordance with (1.3). We have $Z \in \mathcal{L}_{\geqslant 1}$ and the first part of (2.10) says that $\nabla_\lambda F^\bullet_\lambda = 0$, hence $[X_0, Z] = 0$ by Proposition 3.9(i).

Let $Y := G^\bullet_\lambda B_{[\bullet]}$, in accordance with (1.4). We have $Y \in \mathcal{L}_{\geqslant 1}$ and the second part of (2.10) can be rewritten

$$-\mathrm{e}^{-G^\bullet_\lambda} \times \big(\nabla_\lambda(\mathrm{e}^{G^\bullet_\lambda})\big) + \mathrm{e}^{-G^\bullet_\lambda} \times I^\bullet \times \mathrm{e}^{G^\bullet_\lambda} = F^\bullet_\lambda.$$

Let us apply the map $M^\bullet \mapsto M^\bullet B_{[\bullet]}$ to both sides: because of Proposition 3.8(ii) and Proposition 3.9(ii), the image of the left-hand side is $\mathrm{e}^{\operatorname{ad}_Y} X_0 - X_0 + \mathrm{e}^{\operatorname{ad}_Y}(I^\bullet B_{[\bullet]})$, while the image of the right-hand side is $Z$, we thus get

$$\mathrm{e}^{\operatorname{ad}_Y}\big(X_0 + I^\bullet B_{[\bullet]}\big) = X_0 + Z,$$

which is the desired result, since $I^\bullet B_{[\bullet]} = \sum_{n\in\mathcal{N}} B_n$ by (2.7).

3.5. **Proof of the formulas** (1.9)–(1.10) **of Remark 1.3.**

We keep the same assumptions and notations as in Section 3.4.

Let us denote by $\mathcal{E} := \operatorname{End}_{\mathbf{k}}(\mathcal{L})$ the associative algebra consisting of all $\mathbf{k}$-linear operators of the vector space underlying $\mathcal{L}$ (multiplication being defined as operator composition), and by $\mathcal{D}$ the subset of all derivations of the Lie algebra $\mathcal{L}$, which is in fact a Lie subalgebra of $\operatorname{Lie}(\mathcal{E})$ (Lie bracket being defined as operator commutator). For each $m \in \mathbb{N}$, we set

$$\mathcal{E}_{\geqslant m} := \{\, T \in \mathcal{E} \mid T(\mathcal{L}_{\geqslant p}) \subset \mathcal{L}_{\geqslant p+m} \text{ for each } p \in \mathbb{N} \,\}, \quad \mathcal{D}_{\geqslant m} := \mathcal{D} \cap \mathcal{E}_{\geqslant m}. \tag{3.17}$$

It is easy to check that $\mathcal{E}_{\geqslant 0} \supset \mathcal{E}_{\geqslant 1} \supset \mathcal{E}_{\geqslant 2} \supset \ldots$ is a complete filtered associative algebra and $\mathcal{D}_{\geqslant 0} \supset \mathcal{D}_{\geqslant 1} \supset \mathcal{D}_{\geqslant 2} \supset \ldots$ is a complete filtered Lie algebra. Moreover, $\operatorname{ad}\colon \mathcal{L} \to \mathcal{D}_{\geqslant 0}$ is a Lie algebra morphism compatible with the filtrations, in the sense that it maps $\mathcal{L}_{\geqslant m}$ to $\mathcal{D}_{\geqslant m}$. Thus,

$(\mathrm{ad}_{B_n})_{n\in\mathbb{N}}$ is a formally summable family contained in $\mathcal{D}_{\geqslant 1}$ and we are in the situation described at the end of Remark 3.10: with the notation $T_n := \mathrm{ad}_{B_n}$, we may consider the corresponding associative comould and Lie comould, defined by

$$T_{\underline{n}} := \mathrm{ad}_{B_{n_r}} \cdots \mathrm{ad}_{B_{n_1}} \in \mathcal{E}_{\geqslant r}, \qquad T_{[\underline{n}]} := [\mathrm{ad}_{B_{n_r}}, [\ldots [\mathrm{ad}_{B_{n_2}}, \mathrm{ad}_{B_{n_1}}] \ldots]] = \mathrm{ad}_{B_{[\underline{n}]}} \in \mathcal{D}_{\geqslant r}$$

for any $\underline{n} = n_1 \cdots n_r \in \underline{\mathcal{N}}$ (the identity $T_{[\underline{n}]} = \mathrm{ad}_{B_{[\underline{n}]}}$ is due to the Lie algebra morphism property). It follows that $\mathrm{ad}_{M^\bullet B_{[\bullet]}} = M^\bullet T_{[\bullet]}$ for any $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$ and, in the case of the alternal mould $G^\bullet_\lambda$,

$$\mathrm{ad}_Y = \mathrm{ad}_{G^\bullet_\lambda B_{[\bullet]}} = G^\bullet_\lambda T_{[\bullet]} = G^\bullet_\lambda T_\bullet$$

because the restrictions to $\mathrm{Alt}^\bullet(\mathcal{N})$ of the maps $M^\bullet \mapsto M^\bullet T_\bullet$ and $M^\bullet \mapsto M^\bullet T_{[\bullet]}$ coincide. This is (1.9). Remark 3.10 also says that

$$\mathrm{e}^{G^\bullet_\lambda T_{[\bullet]}} = \big(\mathrm{e}^{G^\bullet_\lambda}\big) T_\bullet$$

and, setting $S^\bullet_\lambda := \lambda^*(\mathrm{e}^{G^\bullet}) = \mathrm{e}^{G^\bullet_\lambda}$ (recall that $\lambda^* \colon \mathbf{k}^{\underline{\mathbf{k}}} \to \mathbf{k}^{\underline{\mathcal{N}}}$ is a morphism of associative algebras), we get $\mathrm{e}^{\mathrm{ad}_Y} = S^\bullet_\lambda T_\bullet$, which is (1.10).

### 3.6. **Proof of the addendum to Theorem A.**

We keep the same assumptions and notations as in Section 3.4, except that now $F^\bullet, G^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})$ are moulds satisfying (2.10) (e.g. the ones denoted by $\lambda^*(F^\bullet)$ and $\lambda^*(G^\bullet)$ in Section 3.4).

Let $m \in \mathbb{N}^*$. The set

$$\mathcal{N}_m := \{\, n \in \mathcal{N} \mid \mathrm{ord}(B_n) < m \,\}$$

is finite, as a consequence of the formal summability of the family $(B_n)_{n\in\mathcal{L}}$. We can thus define a "truncation map" $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}} \mapsto M^\bullet_{<m} \in \mathbf{k}^{(\underline{\mathcal{N}})}$ by the formula

$$M^{\varnothing}_{<m} := M^{\varnothing}, \qquad M^{\underline{n}}_{<m} := \mathbb{1}_{\{r<m\}}\, \mathbb{1}_{\{n_1,\ldots,n_r \in \mathcal{N}_m\}}\, M^{\underline{n}} \quad \text{for any nonempty word } \underline{n} = n_1 \cdots n_r \in \underline{\mathcal{N}}$$

and, in our current notations, the formulas (1.11)–(1.12) become

$$Z_m := \sum_{r=1}^{m-1} \sum_{n_1,\ldots,n_r \in \mathcal{N}_m} \frac{1}{r} F^{n_1,\ldots,n_r} B_{[\underline{n}]} = F^\bullet_{<m} B_{[\bullet]}$$

$$Y_m := \sum_{r=1}^{m-1} \sum_{n_1,\ldots,n_r \in \mathcal{N}_m} \frac{1}{r} G^{n_1\ldots,n_r} B_{[\underline{n}]} = G^\bullet_{<m} B_{[\bullet]}.$$

Clearly $\nabla_\lambda F^\bullet = 0$ entails $\nabla_\lambda F^\bullet_{<m} = 0$, hence $[X_0, Z_m] = 0$ by Proposition 3.6. It only remains to be proved that

$$W_m := \mathrm{e}^{\mathrm{ad}_{Y_m}} \Big( X_0 + \sum_{n\in\mathcal{N}} B_n \Big) - X_0 - Z_m$$

has order $\geqslant m$.

**Lemma 3.11.** *If $M^\bullet \in \mathrm{Alt}^\bullet(\mathcal{N})$, then $M^\bullet_{<m} \in \mathrm{Alt}^\bullet_f(\mathcal{N})$.*

*Proof.* Let $\underline{a}$ and $\underline{b}$ be nonempty words and consider the expression $\sum_{\underline{n}\in\underline{\mathcal{N}}} \mathrm{sh}\big(\begin{smallmatrix}\underline{a},\underline{b}\\ \underline{n}\end{smallmatrix}\big) M^{\underline{n}}_{<m}$. We find 0 if $r(\underline{a}\,\underline{b}) \geqslant m$ or if one of the letters of $\underline{a}$ or $\underline{b}$ is outside $\mathcal{N}_m$ (because, then, $\underline{n}$ has the same property whenever $\mathrm{sh}\big(\begin{smallmatrix}\underline{a},\underline{b}\\ \underline{n}\end{smallmatrix}\big) \neq 0$); otherwise we find $\sum_{\underline{n}\in\underline{\mathcal{N}}} \mathrm{sh}\big(\begin{smallmatrix}\underline{a},\underline{b}\\ \underline{n}\end{smallmatrix}\big) M^{\underline{n}}$, which is also 0 if $M^\bullet$ is supposed to be alternal. $\square$

Hence $F^\bullet_{<m}$ and $G^\bullet_{<m}$ are alternal and we can use Proposition 3.8(ii) and Proposition 3.9(ii) with $Y_m = G^\bullet_{<m} B_{[\bullet]}$ to rewrite $W_m = \mathrm{e}^{\mathrm{ad}_{Y_m}} X_0 - X_0 + \mathrm{e}^{\mathrm{ad}_{Y_m}}(I^\bullet B_{[\bullet]}) - F^\bullet_{<m} B_{[\bullet]}$ as

$$W_m = \big(\mathrm{e}^{-G^\bullet_{<m}} \times E^\bullet\big) B_{[\bullet]}, \qquad E^\bullet := -\nabla_\lambda(\mathrm{e}^{G^\bullet_{<m}}) + I^\bullet \times \mathrm{e}^{G^\bullet_{<m}} - \mathrm{e}^{G^\bullet_{<m}} \times F^\bullet_{<m}. \tag{3.18}$$

Let $C^\bullet := F^\bullet - F^\bullet_{<m}$, $\tilde{C}^\bullet := G^\bullet - G^\bullet_{<m}$ and $D^\bullet := \mathrm{e}^{G^\bullet} - \mathrm{e}^{G^\bullet_{<m}}$. Since $-\nabla_\lambda\big(\mathrm{e}^{G^\bullet}\big) + I^\bullet \times \mathrm{e}^{G^\bullet} - \mathrm{e}^{G^\bullet} \times F^\bullet = 0$, we get

$$E^\bullet = \nabla_\lambda D^\bullet - I^\bullet \times D^\bullet + D^\bullet \times F^\bullet + \mathrm{e}^{G^\bullet_{<m}} \times C^\bullet. \tag{3.19}$$

**Lemma 3.12.** *(i) Suppose $M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$ and $M^\bullet_{<m} = 0$. Then $M^\bullet B_{[\bullet]} \in \mathcal{L}_{\geqslant m}$.*
*(ii) Suppose $M^\bullet, N^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$ and $M^\bullet_{<m} = 0$. Then $(M^\bullet \times N^\bullet)_{<m} = (N^\bullet \times M^\bullet)_{<m} = 0$.*

*Proof.* Suppose $M^\bullet_{<m} = 0$.
(i) For any word $\underline{n} = n_1 \cdots n_r$, $M^{\underline{n}} \neq 0$ implies $\max\{r, \mathrm{ord}(B_{n_1}), \ldots, \mathrm{ord}(B_{n_r})\} \geqslant m$, but $\mathrm{ord}(B_{[\underline{n}]}) \geqslant \max\{r, \mathrm{ord}(B_{n_1}), \ldots, \mathrm{ord}(B_{n_r})\}$, hence $\mathrm{ord}(M^{\underline{n}} B_{[\underline{n}]}) \geqslant m$ in all cases.
(ii) Suppose $\underline{n} = n_1 \cdots n_r$ with $r < m$ and $n_1, \ldots, n_r \in \mathcal{N}_m$. We have $(M^\bullet \times N^\bullet)^{\underline{n}} = \sum M^{\underline{a}} N^{\underline{b}}$ with summation over all pairs of words such that $\underline{a}\,\underline{b} = \underline{n}$, which entails $M^{\underline{a}} = 0$ in each term of the sum, and similarly for $N^\bullet \times M^\bullet$. $\square$

We have $C^\bullet_{<m} = \tilde{C}^\bullet_{<m} = 0$, and $D^\bullet = \sum_{k \geqslant 0} \frac{1}{(k+1)!}\big((G^\bullet)^{\times(k+1)} - (G^\bullet_{<m})^{\times(k+1)}\big)$ with

$$(G^\bullet)^{\times(k+1)} - (G^\bullet_{<m})^{\times(k+1)} = \sum_{k=p+q} (G^\bullet_{<m})^{\times p} \times \tilde{C}^\bullet \times (G^\bullet)^{\times q} \quad \text{for each } k \geqslant 0,$$

whence $D^\bullet_{<m} = 0$ by Lemma 3.12(ii). In view of (3.19), it follows, again by Lemma 3.12(ii), that $\big(\mathrm{e}^{-G^\bullet_{<m}} \times E^\bullet\big)_{<m} = 0$, whence $W_m \in \mathcal{L}_{\geqslant m}$ by (3.18) and Lemma 3.12(i).

## 4. Resolution of the mould equation and proof of Theorem B

With the view of proving Theorem B, we now give ourselves a nonempty set $\mathcal{N}$, a field $\mathbf{k}$ of characteristic zero and a map $\lambda\colon \mathcal{N} \to \mathbf{k}$.

Part (i) of the statement of Theorem B requires that, for each $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, we prove the existence and uniqueness of a pair $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ solving (2.10)–(2.11). As explained in Section 2.5.5, with the change of unknown mould $S^\bullet := \mathrm{e}^{G^\bullet}$, this is equivalent to proving the existence and uniqueness of a pair $(S^\bullet, F^\bullet) \in \mathrm{Sym}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ solving equation (2.20) and satisfying

$$\left[{}^{\mathrm{inv}}S^\bullet \times \nabla_1 S^\bullet\right]_{\lambda=0} = A^\bullet. \tag{4.1}$$

Heuristically, here is what happens: it is easy to see that, apart from the exceptional case in which $\lambda(\underline{n}) \neq 0$ for every nonempty word $\underline{n}$ (in which case $\mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N}) = \{0\}$ and there is a unique solution $(S^\bullet, F^\bullet)$ to (2.20) in $\mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ such that $S^\varnothing = 1$), equation (2.20) has in general infinitely many solutions $(S^\bullet, F^\bullet) \in \mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ such that $S^\varnothing = 1$ (because one is free to assign an arbitrary value to $S^{\underline{n}}$ whenever $\lambda(\underline{n}) = 0$), but what is not obvious is the existence of at least one solution *with $S^\bullet$ symmetral and $F^\bullet$ alternal*; adding the requirement (4.1) removes the freedom: then we get a unique solution $(S^\bullet, F^\bullet)$ in $\mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ such that $S^\varnothing = 1$, and we are left with the problem of proving that this solution is in $\mathrm{Sym}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$. This will follow from the alternality of $A^\bullet$ at the price of an excursion in the space of "dimoulds".

### 4.1. **The associative algebra of dimoulds.**

The material in this section is essentially taken from [Sau09].

We call *dimould* any map $\underline{\mathcal{N}} \times \underline{\mathcal{N}} \to \mathbf{k}$. We denote by $M^{\bullet,\bullet}$ the dimould whose value on a pair of words $(\underline{a}, \underline{b})$ is $M^{\underline{a},\underline{b}}$. The set $\mathbf{k}^{\underline{\mathcal{N}} \times \underline{\mathcal{N}}}$ of all dimoulds is clearly a linear space over $\mathbf{k}$, it is also an associative $\mathbf{k}$-algebra for the *dimould multiplication* $(M^{\bullet,\bullet}, N^{\bullet,\bullet}) \mapsto P^{\bullet,\bullet} = M^{\bullet,\bullet} \times N^{\bullet,\bullet}$ defined by a formula analogous to (2.1):

$$P^{\underline{a},\underline{b}} := \sum_{(\underline{a},\underline{b}) = (\underline{a}^1,\underline{b}^1)(\underline{a}^2,\underline{b}^2)} M^{(\underline{a}^1,\underline{b}^1)} N^{(\underline{a}^2,\underline{b}^2)},$$

where the concatenation in $\underline{\mathcal{N}} \times \underline{\mathcal{N}}$ is defined by $(\underline{a}^1, \underline{b}^1)(\underline{a}^2, \underline{b}^2) = (\underline{a}^1\,\underline{a}^2, \underline{b}^1\,\underline{b}^2)$.

Examples of dimoulds are the *decomposable dimoulds*, namely the dimoulds of the form

$$P^{\bullet,\bullet} = M^\bullet \otimes N^\bullet,$$

where it is meant that $M^\bullet$ and $N^\bullet$ are (ordinary) moulds and $P^{\underline{a},\underline{b}} = M^{\underline{a}} N^{\underline{b}}$. Note that

$$(M_1^\bullet \otimes N_1^\bullet) \times (M_2^\bullet \otimes N_2^\bullet) = (M_1^\bullet \times M_2^\bullet) \otimes (N_1^\bullet \times N_2^\bullet) \tag{4.2}$$

for any four moulds $M_1^\bullet, N_1^\bullet, M_2^\bullet, N_2^\bullet$.

Using the shuffling coefficients of Definition 2.2, we define a linear map

$$\Delta\colon\ M^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}} \mapsto P^{\bullet,\bullet} = \Delta(M^\bullet) \in \mathbf{k}^{\underline{\mathcal{N}} \times \underline{\mathcal{N}}} \tag{4.3}$$

as follows:

$$P^{\underline{a},\underline{b}} := \sum_{\underline{n} \in \underline{\mathcal{N}}} \mathrm{sh}\big(\begin{smallmatrix}\underline{a},\,\underline{b}\\ \underline{n}\end{smallmatrix}\big) M^{\underline{n}} \quad \text{for any } (\underline{a}, \underline{b}) \in \underline{\mathcal{N}} \times \underline{\mathcal{N}}. \tag{4.4}$$

We thus can rephrase the definition of alternality given in Definition 2.2 and the definition of symmetrality given in (3.16):

$$\text{A mould } M^\bullet \text{ is alternal if and only if } \Delta(M^\bullet) = M^\bullet \otimes 1^\bullet + 1^\bullet \otimes M^\bullet. \tag{4.5}$$

$$\text{It is symmetral if and only if } M^\varnothing = 1 \text{ and } \Delta(M^\bullet) = M^\bullet \otimes M^\bullet. \tag{4.6}$$

It is proved in [Sau09, Sec. 5.2] that[5]

$$\Delta\colon \mathbf{k}^{\underline{\mathcal{N}}} \to \mathbf{k}^{\underline{\mathcal{N}}\times\underline{\mathcal{N}}} \text{ is an associative algebra morphism.} \tag{4.7}$$

We end this section with an example of *dimould derivation*, i.e. a derivation of the dimould algebra $\mathbf{k}^{\underline{\mathcal{N}}\times\underline{\mathcal{N}}}$.

**Lemma 4.1.** *Let $\varphi\colon \mathcal{N} \to \mathbf{k}$ denote an abitrary function, extended to $\underline{\mathcal{N}}$ by (2.8). Then the formula*

$$\widetilde{\nabla}_\varphi\colon P^{\bullet,\bullet} \mapsto Q^{\bullet,\bullet}, \qquad Q^{\underline{a},\underline{b}} := \big(\varphi(\underline{a}) + \varphi(\underline{b})\big) P^{\underline{a},\underline{b}} \quad \text{for all } \underline{a}, \underline{b} \in \underline{\mathcal{N}}$$

*defines a $\mathbf{k}$-linear operator $\widetilde{\nabla}_\varphi$ of $\mathbf{k}^{\underline{\mathcal{N}}\times\underline{\mathcal{N}}}$ which is a dimould derivation and satisfies*

$$\widetilde{\nabla}_\varphi(M^\bullet \otimes N^\bullet) = (\nabla_\varphi M^\bullet) \otimes N^\bullet + M^\bullet \otimes \nabla_\varphi N^\bullet \tag{4.8}$$

$$\Delta(\nabla_\varphi M^\bullet) = \widetilde{\nabla}_\varphi \Delta(M^\bullet) \tag{4.9}$$

*for any two moulds $M^\bullet$ and $N^\bullet$, where $\nabla_\varphi$ is the mould derivation defined by (2.9).*

The proof of Lemma 4.1 is left to the reader (use $\operatorname{sh}\big(\genfrac{}{}{}{}{\underline{a},\underline{b}}{\underline{n}}\big) \neq 0 \ \Rightarrow\ \varphi(\underline{a}) + \varphi(\underline{b}) = \varphi(\underline{n})$ for the last property).

### 4.2. **Proof of Part (i) of Theorem B.**

Let $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$. As explained at the beginning of Section 4, the strategy is first to check the existence and uniqueness of a pair of moulds $(S^\bullet, F^\bullet) \in \mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ satisfying (2.20) and (4.1) and $S^\varnothing = 1$, and then to prove (with the help of dimoulds) that $(S^\bullet, F^\bullet) \in \mathrm{Sym}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$.

*4.2.1.* Let us introduce an extra unknown mould $N^\bullet = {}^{\mathrm{inv}}S^\bullet \times \nabla_1 S^\bullet$, so that finding a solution $(S^\bullet, F^\bullet)$ to (2.20) and (4.1) is equivalent to finding a solution $(S^\bullet, F^\bullet, N^\bullet)$ to the system of

[5] In this paper we have denoted by $\Delta$ the map which was denoted by $\tau$ in [Sau09], because this map is essentially the coproduct of a Hopf algebra structure that one can define and the notation $\Delta$ is more common for coproducts.

equations

$$\nabla_\lambda S^\bullet = I^\bullet \times S^\bullet - S^\bullet \times F^\bullet \tag{4.10}$$

$$\nabla_1 S^\bullet = S^\bullet \times N^\bullet \tag{4.11}$$

$$\nabla_\lambda F^\bullet = 0 \tag{4.12}$$

$$N^\bullet_{\lambda=0} = A^\bullet. \tag{4.13}$$

The system (4.10)–(4.13), in presence of the condition $S^\varnothing = 1$, amounts to $F^\varnothing = N^\varnothing = 0$ and, for each nonempty word $\underline{n}$,

$$\lambda(\underline{n})\, S^{\underline{n}} + F^{\underline{n}} = S^{`\underline{n}} - \sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, F^{\underline{b}} \tag{4.14}$$

$$r(\underline{n})\, S^{\underline{n}} - N^{\underline{n}} = \sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, N^{\underline{b}} \tag{4.15}$$

$$\lambda(\underline{n}) \neq 0 \ \Rightarrow\ F^{\underline{n}} = 0 \tag{4.16}$$

$$\lambda(\underline{n}) = 0 \ \Rightarrow\ N^{\underline{n}} = A^{\underline{n}} \tag{4.17}$$

with $`\underline{n}$ denoting the word $\underline{n}$ deprived from its first letter.

We thus find a unique solution by induction on $r(\underline{n})$: we must take $S^\varnothing = 1$, $F^\varnothing = N^\varnothing = 0$ and, for $r(\underline{n}) \geqslant 1$,

$$\lambda(\underline{n}) \neq 0 \ \Rightarrow\ \begin{cases} F^{\underline{n}} = 0 \\[2ex] S^{\underline{n}} = \dfrac{1}{\lambda(\underline{n})}\Big(S^{`\underline{n}} - \displaystyle\sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, F^{\underline{b}}\Big) \\[2ex] N^{\underline{n}} = r(\underline{n})\, S^{\underline{n}} - \displaystyle\sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, N^{\underline{b}} \end{cases} \tag{4.18}$$

$$\lambda(\underline{n}) = 0 \ \Rightarrow\ \begin{cases} F^{\underline{n}} = S^{`\underline{n}} - \displaystyle\sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, F^{\underline{b}} \\[2ex] N^{\underline{n}} = A^{\underline{n}} \\[2ex] S^{\underline{n}} = \dfrac{1}{r(\underline{n})}\Big(A^{\underline{n}} + \displaystyle\sum_{\underline{n}=\underline{a}\,\underline{b},\ \underline{a},\underline{b}\neq\varnothing} S^{\underline{a}}\, N^{\underline{b}}\Big). \end{cases} \tag{4.19}$$

*4.2.2.* We now check that, in the unique solution constructed above, $S^\bullet$ is symmetral and $F^\bullet$ is alternal. Making use of the dimould formalism of Section 4.1, and in particular of the associative algebra morphism $\Delta$ defined by (4.3)–(4.4), we set

$$A^{\bullet,\bullet} := \Delta(A^\bullet), \quad S^{\bullet,\bullet} := \Delta(S^\bullet), \quad F^{\bullet,\bullet} := \Delta(F^\bullet), \quad N^{\bullet,\bullet} := \Delta(N^\bullet).$$

Our assumption amounts to $A^{\bullet,\bullet} = A^\bullet \otimes 1^\bullet + 1^\bullet \otimes A^\bullet$ and we are to prove $S^{\bullet,\bullet} = S^\bullet \otimes S^\bullet$ and $F^{\bullet,\bullet} = F^\bullet \otimes 1^\bullet + 1^\bullet \otimes F^\bullet$. Note that $S^{\varnothing,\varnothing} = S^\varnothing = 1$.

In view of Lemma 4.1, the dimould derivations $\widetilde\nabla_\lambda$ and $\widetilde\nabla_1$ are defined by

$$\widetilde\nabla_\lambda M^{\underline{a},\underline{b}} := \big(\lambda(\underline{a}) + \lambda(\underline{b})\big) M^{\underline{a},\underline{b}} \quad \text{and} \quad \widetilde\nabla_1 M^{\underline{a},\underline{b}} := \big(r(\underline{a}) + r(\underline{b})\big) M^{\underline{a},\underline{b}} \quad \text{for all } \underline{a},\underline{b} \in \underline{\mathcal{N}}$$

for any dimould $M^{\bullet,\bullet}$. Applying $\Delta$ to each equation of the system (4.10)–(4.13), we get

$$\widetilde\nabla_\lambda S^{\bullet,\bullet} = \Delta(I^\bullet) \times S^{\bullet,\bullet} - S^{\bullet,\bullet} \times F^{\bullet,\bullet} \tag{4.20}$$

$$\widetilde\nabla_1 S^{\bullet,\bullet} = S^{\bullet,\bullet} \times N^{\bullet,\bullet} \tag{4.21}$$

$$\widetilde\nabla_\lambda F^{\bullet,\bullet} = 0 \tag{4.22}$$

$$N^{\bullet,\bullet}_{\lambda=0} = A^{\bullet,\bullet}. \tag{4.23}$$

Here we have used the associative algebra morphism property (4.7) of $\Delta$ and the identity (4.9) with $\nabla_\lambda$ and $\nabla_1$; moreover, we have denoted by $N^{\bullet,\bullet}_{\lambda=0}$ the resonant part of the dimould $N^{\bullet,\bullet}$ defined by

$$N^{\underline{a},\underline{b}}_{\lambda=0} := \mathbb{1}_{\{\lambda(\underline{a})+\lambda(\underline{b})=0\}}\, N^{\underline{a},\underline{b}} \qquad \text{for any } (\underline{a},\underline{b}) \in \underline{\mathcal{N}} \times \underline{\mathcal{N}}$$

and used the obvious identity $\big(\Delta(N^\bullet)\big)_{\lambda=0} = \Delta(N^\bullet_{\lambda=0})$ (due to the fact that $\mathrm{sh}\big(\begin{smallmatrix}\underline{a},\underline{b}\\ \underline{n}\end{smallmatrix}\big) \neq 0 \;\Rightarrow\; \lambda(\underline{a}) + \lambda(\underline{b}) = \lambda(\underline{n})$).

We now observe that the system of dimould equations (4.20)–(4.23) has a unique solution $(S^{\bullet,\bullet}, F^{\bullet,\bullet}, N^{\bullet,\bullet})$ such that $S^{\varnothing,\varnothing} = 1$. Indeed, these equations entail $F^{\varnothing,\varnothing} = N^{\varnothing,\varnothing} = 0$ and, by evaluating them on a pair of words $(\underline{a},\underline{b}) \neq (\varnothing,\varnothing)$, we get equations analogous to (4.14)–(4.17) which allow to determine $S^{\underline{a},\underline{b}}$, $F^{\underline{a},\underline{b}}$ and $N^{\underline{a},\underline{b}}$ by induction on $r(\underline{a}) + r(\underline{b})$ (distinguishing the cases $\lambda(\underline{a}) + \lambda(\underline{b}) = 0$ or $\neq 0$).

Since $\Delta(I^\bullet) = I^\bullet \otimes 1^\bullet + 1^\bullet \otimes I^\bullet$ and $A^{\bullet,\bullet} = A^\bullet \otimes 1^\bullet + 1^\bullet \otimes A^\bullet$, it is easy to check directly that $(S^\bullet \otimes S^\bullet, F^\bullet \otimes 1^\bullet + 1^\bullet \otimes F^\bullet, N^\bullet \otimes 1^\bullet + 1^\bullet \otimes N^\bullet)$ is a solution of the system (4.20)–(4.23) with the initial condition $(S^\bullet \otimes S^\bullet)^{\varnothing,\varnothing} = 1$ (one just has to use (4.2), (4.8), (4.10)–(4.13) and the identities $(N^\bullet \otimes 1^\bullet)_{\lambda=0} = N^\bullet_{\lambda=0} \otimes 1^\bullet$, $(1^\bullet \otimes N^\bullet)_{\lambda=0} = 1^\bullet \otimes N^\bullet_{\lambda=0}$).

The uniqueness of the solution of the system of dimould equations implies

$$(S^{\bullet,\bullet}, F^{\bullet,\bullet}, N^{\bullet,\bullet}) = (S^\bullet \otimes S^\bullet, F^\bullet \otimes 1^\bullet + 1^\bullet \otimes F^\bullet, N^\bullet \otimes 1^\bullet + 1^\bullet \otimes N^\bullet)$$

in particular $S^\bullet$ is symmetral and $F^\bullet$ is alternal.

*4.2.3.* The induction formulas (4.18)–(4.19) that we have obtained for $F^\bullet$ and $S^\bullet$ coincide with (2.15)–(2.16). Setting $G^\bullet = \log S^\bullet$, we get an alternal mould, inductively determined by (2.17).

This ends the proof of Part (i) of Theorem B.

4.3. **Proof of Part (ii) of Theorem B.**

*4.3.1.* Recall that the mould exponential $G^\bullet \mapsto S^\bullet = \mathrm{e}^{G^\bullet}$ is a bijection between the set of all moulds $G^\bullet$ such that $G^\varnothing = 0$ and the set of all moulds $S^\bullet$ such that $S^\varnothing = 1$, which induces a bijection $\mathrm{Alt}^\bullet(\mathcal{N}) \to \mathrm{Sym}^\bullet(\mathcal{N})$. There is thus a bijection between the solutions $(F^\bullet, G^\bullet) \in \mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ to equation (2.10) such that $G^\varnothing = 0$ and the solutions $(F^\bullet, S^\bullet) \in \mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\mathcal{N}}}$ to equation (2.20) such that $S^\varnothing = 1$. We rewrite equation (2.20) as

$$F^\bullet = {}^{\mathrm{inv}}S^\bullet \times I^\bullet \times S^\bullet - {}^{\mathrm{inv}}S^\bullet \times \nabla_\lambda S^\bullet, \tag{4.24}$$

$$\nabla_\lambda F^\bullet = 0. \tag{4.25}$$

Starting with a solution $(F^\bullet, G^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Alt}^\bullet(\mathcal{N})$ to (2.10) and setting $S^\bullet := \mathrm{e}^{G^\bullet} \in \mathrm{Sym}^\bullet(\mathcal{N})$, we get a solution $(F^\bullet, S^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ to (4.24)–(4.25); using the change $K^\bullet = \mathrm{e}^{J^\bullet}$ (as in Section 2.5.7), we are asked to prove that the map

$$K^\bullet \mapsto (\tilde{F}^\bullet, \tilde{S}^\bullet) = \big({}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet,\ S^\bullet \times K^\bullet\big) \tag{4.26}$$

establishes a one-to-one correspondence between $\mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$ and the set of all solutions $(\tilde{F}^\bullet, \tilde{S}^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ to (4.24)–(4.25), and that

$$\Big[{}^{\mathrm{inv}}\tilde{S}^\bullet \times \nabla_1 \tilde{S}^\bullet\Big]_{\lambda=0} = {}^{\mathrm{inv}}K^\bullet \times \mathscr{I}_\lambda(G^\bullet) \times K^\bullet + {}^{\mathrm{inv}}K^\bullet \times \nabla_1 K. \tag{4.27}$$

*4.3.2.* Suppose that $K^\bullet \in \mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$ and define $(\tilde{F}^\bullet, \tilde{S}^\bullet)$ by (4.26). Since $\tilde{S}^\bullet = S^\bullet \times K^\bullet$, this mould is symmetral (recall that $(\mathrm{Sym}^\bullet(\mathcal{N}), \times)$ is a group—see e.g. [Sau09, Prop. 5.1]); since $\nabla_\lambda$ is a derivation which annihilates $K^\bullet$, we have $\nabla_\lambda \tilde{S}^\bullet = (\nabla_\lambda S^\bullet) \times K^\bullet$ and

$${}^{\mathrm{inv}}\tilde{S}^\bullet \times I^\bullet \times \tilde{S}^\bullet - {}^{\mathrm{inv}}\tilde{S} \times \nabla_\lambda \tilde{S} = {}^{\mathrm{inv}}K^\bullet \times \big({}^{\mathrm{inv}}S^\bullet \times I^\bullet \times S^\bullet - {}^{\mathrm{inv}}S \times \nabla_\lambda S\big) \times K^\bullet,$$

which, by (4.24), is ${}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet = \tilde{F}^\bullet$. Thus, $(\tilde{F}^\bullet, \tilde{S}^\bullet)$ satisfies (4.24).

On the other hand, by (2.19) and Proposition 3.8(ii), $\tilde{F} = {}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet$ is alternal. It is easy to check that $\tilde{F}^\bullet$ satisfies (4.25) because $F^\bullet$ satisfies (4.25): $0 = {}^{\mathrm{inv}}K^\bullet \times \nabla_\lambda F^\bullet \times K^\bullet = \nabla_\lambda \tilde{F}$. It is so because $\nabla_\lambda$ is derivation which annihilates both $K^\bullet$ and ${}^{\mathrm{inv}}K^\bullet$; the fact that also ${}^{\mathrm{inv}}K^\bullet$ is $\lambda$-resonant (i.e. $\nabla_\lambda {}^{\mathrm{inv}}K^\bullet = 0$) is an elementary property of $\lambda$-resonant moulds, which is part of

**Lemma 4.2.** *Suppose that $M^\bullet$ is a $\lambda$-resonant mould. Then also $\nabla_1 M^\bullet$ is $\lambda$-resonant, and*

$$[M^\bullet \times N^\bullet]_{\lambda=0} = M^\bullet \times N^\bullet_{\lambda=0}, \quad [N^\bullet \times M^\bullet]_{\lambda=0} = N^\bullet_{\lambda=0} \times M^\bullet \qquad \textit{for any mould } N^\bullet.$$

*If moreover $M^\bullet$ is invertible, then also ${}^{\mathrm{inv}}M^\bullet$ is $\lambda$-resonant.*

The proof of Lemma 4.2 is left to the reader.

We now compute the gauge generator of $\log \tilde{S}^\bullet$: by Lemma 4.2, the $\lambda$-resonant part of

$$^{\mathrm{inv}}\tilde{S}^\bullet \times \nabla_1 \tilde{S}^\bullet = {}^{\mathrm{inv}}K^\bullet \times {}^{\mathrm{inv}}S^\bullet \times \big((\nabla_1 S^\bullet) \times K^\bullet + S^\bullet \times \nabla_1 K^\bullet\big)$$

is ${}^{\mathrm{inv}}K^\bullet \times \big[{}^{\mathrm{inv}}S^\bullet \times \nabla_1 S^\bullet\big]_{\lambda=0} \times K^\bullet + {}^{\mathrm{inv}}K^\bullet \times \nabla_1 K^\bullet = {}^{\mathrm{inv}}K^\bullet \times \mathscr{I}_\lambda(G^\bullet) \times K^\bullet + {}^{\mathrm{inv}}K^\bullet \times \nabla_1 K^\bullet$. This is (4.27).

*4.3.3.* Conversely, suppose that $(\tilde{F}^\bullet, \tilde{S}^\bullet) \in \mathrm{Alt}^\bullet(\mathcal{N}) \times \mathrm{Sym}^\bullet(\mathcal{N})$ is a solution to (4.24)–(4.25). We define $K^\bullet := {}^{\mathrm{inv}}S^\bullet \times \tilde{S}^\bullet \in \mathrm{Sym}^\bullet(\mathcal{N})$. Inserting

$$\tilde{S}^\bullet = S^\bullet \times K^\bullet \tag{4.28}$$

in $\tilde{F}^\bullet = {}^{\mathrm{inv}}\tilde{S}^\bullet \times I^\bullet \times \tilde{S}^\bullet - {}^{\mathrm{inv}}\tilde{S}^\bullet \times \nabla_\lambda \tilde{S}^\bullet$, we get

$$\tilde{F}^\bullet = {}^{\mathrm{inv}}K^\bullet \times \big({}^{\mathrm{inv}}S^\bullet \times I^\bullet \times S^\bullet \times K^\bullet - {}^{\mathrm{inv}}S^\bullet \times \nabla_\lambda(S^\bullet \times K^\bullet)\big) = {}^{\mathrm{inv}}K^\bullet \times (F^\bullet \times K^\bullet - \nabla_\lambda K^\bullet), \tag{4.29}$$

i.e. $\nabla_\lambda K^\bullet = F^\bullet \times K^\bullet - K^\bullet \times \tilde{F}^\bullet$. We are in a position to apply

**Lemma 4.3.** *Suppose that $M^\bullet, N^\bullet, P^\bullet \in \mathbf{k}^{\underline{\mathcal{N}}}$, $M^\varnothing = N^\varnothing = 0$, $M^\bullet$ and $N^\bullet$ are $\lambda$-resonant and*

$$\nabla_\lambda P^\bullet = M^\bullet \times P^\bullet - P^\bullet \times N^\bullet. \tag{4.30}$$

*Then $P^\bullet$ is $\lambda$-resonant.*

Taking Lemma 4.3 for granted, we thus obtain that $K^\bullet$ is $\lambda$-resonant, hence $K^\bullet \in \mathrm{Sym}^\bullet_{\lambda=0}(\mathcal{N})$, and (4.29) yields $\tilde{F} = {}^{\mathrm{inv}}K^\bullet \times F^\bullet \times K^\bullet$, which together with (4.28) gives $(\tilde{F}^\bullet, \tilde{S}^\bullet)$ as the image of $K^\bullet$ by the map (4.26). The proof of Theorem B(ii) is then complete.

*Proof of Lemma 4.3.* Let us show that

$$\lambda(\underline{n}) P^{\underline{n}} = 0 \tag{4.31}$$

for every $\underline{n} \in \underline{\mathcal{N}}$ by induction on $r(\underline{n})$. The property holds for $\underline{n} = \varnothing$ or, more generally, for $\lambda(\underline{n}) = 0$, we thus suppose that $\underline{n} \in \underline{\mathcal{N}}$ has $r(\underline{n}) \geqslant 1$ and $\lambda(\underline{n}) \neq 0$, and that (4.31) holds for all words of length $< r(\underline{n})$. It follows from (4.30) that

$$\lambda(\underline{n}) P^{\underline{n}} = \sum_{\underline{n}=\underline{a}\,\underline{b}} (M^{\underline{a}} P^{\underline{b}} - P^{\underline{a}} N^{\underline{b}}) = \sum_{\underline{n}=\underline{a}\,\underline{b}}^{*} (M^{\underline{a}} P^{\underline{b}} - P^{\underline{a}} N^{\underline{b}}), \tag{4.32}$$

where the symbol $\sum^*$ indicates that we can restrict the summation to non-trivial decompositions (it is so because $M^{\underline{n}} = N^{\underline{n}} = 0$, since $\lambda(\underline{n}) \neq 0$, and $M^\varnothing = N^\varnothing = 0$). But, in the right-hand side of (4.32), each term between parentheses vanishes, because either $\lambda(\underline{a}) \neq 0$ and $M^{\underline{a}} = P^{\underline{a}} = 0$ (by the assumption on $M^{\underline{a}}$ and the inductive hypothesis), or $\lambda(\underline{a}) = 0$, but then $\lambda(\underline{b}) \neq 0$ and $M^{\underline{b}} = P^{\underline{b}} = 0$ (for similar reasons). $\square$

# Five dynamical applications

We now turn to examples of application of Theorem A. The Lie algebras in these examples will consist of vector fields with their natural Lie brackets $[\cdot\,,\cdot]_{\rm vf}$ or, in presence of a symplectic structure, Hamiltonian functions with the Lie bracket $[\cdot\,,\cdot]_{\rm ham} := \{\cdot\,,\cdot\}$ (Poisson bracket) or, in the quantum case, operators of a Hilbert space with the Lie bracket $[\cdot\,,\cdot]_{\rm qu} := \frac{1}{{\rm i}\hbar} \times$ commutator. We will deal with formal objects (i.e. defined by means of formal series, either in the dynamical variables or in some external parameter), and this gives rise to a natural Lie algebra filtration.

## 5. **Poincaré-Dulac normal forms**

**5.1** Let $N \in \mathbb{N}^*$. A formal vector field is the same thing as a derivation of the algebra of formal series $\mathbb{C}[[z_1,\dots,z_N]]$ and is of the form

$$X = \sum_{j=1}^{N} v_j(z_1,\dots,z_N)\partial_{z_j}.$$

We take $\mathbf{k} := \mathbb{C}$ and $\mathcal{L} :=$ the Lie algebra of formal vector fields whose components $v_j$ have no constant term. We get a complete filtered algebra by setting $X \in \mathcal{L}_{\geqslant m}$ if its components $v_j$, as formal series, have order $\geqslant m+1$.

Let $X \in \mathcal{L}$. The *formal normalization problem* consists in finding a formal change of variables which simplifies the expression of $X$ as much as possible. We assume that $X$ has a diagonal linear part:

$$X_0 = \sum_{j=1}^{N} \omega_j z_j \partial_{z_j}$$

with "spectrum vector" $\omega = (\omega_1,\dots,\omega_N) \in \mathbb{C}^N$. The components of $B := X - X_0$ have order $\geqslant 2$, hence, introducing

$$\mathcal{M} := \big\{\, (j,k) \in \{1,\dots,N\} \times \mathbb{N}^N \mid |k| \geqslant 2 \,\big\},$$

we can write the expansion of $X - X_0$ as $B = \sum_{(j,k)\in\mathcal{M}} b_{j,k} z^k \partial_{z_j}$ with coefficients $b_{j,k} \in \mathbb{C}$. It turns out that the monomial vector fields $z^k\partial_{z_j}$ are eigenvectors of ${\rm ad}_{X_0}$:

$$\Big[X_0, z^k\partial_{z_j}\Big]_{\rm vf} = \big(\langle k,\omega\rangle - \omega_j\big) z^k \partial_{z_j} \quad \text{for each } (j,k) \in \mathcal{M} \tag{5.1}$$

(where $\langle\cdot\,,\cdot\rangle$ denotes the standard scalar product), we thus set

$$\mathcal{N} := \big\{ \langle k,\omega\rangle - \omega_j \mid (j,k)\in\mathcal{M} \text{ and } b_{j,k}\neq 0 \big\} \subset \mathbb{C},$$

$$B_\lambda := \sum_{\substack{(j,k)\in\mathcal{M}\text{ such that}\\ \langle k,\omega\rangle-\omega_j=\lambda}} b_{j,k} z^k \partial_{z_j} \quad \text{for each } \lambda\in\mathcal{N},$$

so that $X = X_0 + \sum\limits_{\lambda\in\mathcal{N}} B_\lambda$ and $[X_0,B_\lambda]_{\mathrm{vf}} = \lambda B_\lambda$ for each $\lambda\in\mathcal{N}$.

**5.2** Let us apply Theorem A: with each choice of $A^\bullet \in \mathrm{Alt}_0^\bullet(\mathcal{N})$ is associated a pair of alternal moulds, $F^\bullet$ and $G^\bullet$ explicitly given by (2.14)–(2.17), which give rise to formal vector fields $Z$ and $Y$ such that (1.2) holds: the automorphism $\mathrm{e}^{\mathrm{ad}_Y}$ of $\mathcal{L}$ maps $X = X_0 + B$ to $X_0 + Z$ and $[X_0, Z]_{\mathrm{vf}} = 0$. Moreover, $Z$ and $Y$ are explicitly given by the expansions (1.3)–(1.4) (with the convention of Definition 2.4: the map $\lambda$ is to be interpreted as the inclusion map $\mathcal{N}\hookrightarrow\mathbb{C}$).

In this context, a formal vector field which commutes with $X_0$ is called "resonant". According to (5.1), this means that it is a sum of "resonant monomials", i.e. multiples of elementary vector fields of the form $z^k\partial_{z_j}$ with

$$(j,k)\in\mathcal{M} \quad \text{such that} \quad \langle k,\omega\rangle - \omega_j = 0. \tag{5.2}$$

It may happen that there exist no resonant monomial at all: one says that the spectrum vector $\omega$ is "non-resonant" if equation (5.2) has no solution (a kind of arithmetical condition). Necessarily $Z = 0$ in that case (although $F^\bullet$ might be nonzero).

The first part of (1.2) thus says that $Z$ is a formal resonant vector field; classically, $X_0 + Z$ is called a *Poincaré-Dulac normal form.* In [EV95], the particular Poincaré-Dulac normal form corresponding to the choice $A^\bullet = 0$ (zero gauge solution of equation (2.10)) is called "regal prenormal form".

The automorphism $\mathrm{e}^{\mathrm{ad}_Y}$ of $\mathcal{L}$ is nothing but the action of the formal flow $\Phi$ of $Y$ at time 1 by pull-back: $\mathrm{e}^{\mathrm{ad}_Y}X = \Phi_*^{-1}X$, hence the second part of (1.2) says that $\Phi_*^{-1}X = X_0 + Z$, which corresponds to the formal change of coordinates $z \mapsto \Phi^{-1}(z)$ obtained by flowing at time 1 along $-Y$.

We have thus recovered the classical results by Poincaré and Dulac, according to which *one can formally conjugate $X$ to its linear part $X_0$ when $\omega$ is non-resonant and, in the general case, to a formal vector field the expression of which contains only resonant monomials.*

It is well known that, in general, there is more than one Poincaré-Dulac normal form.

**5.3** For a resonant vector $\omega$, there may be only one resonance relation (5.2) (e.g. for $\omega = (2,1)$ in dimension $N = 2$) or infinitely many of them (e.g. for $\omega = (-1,1)$). A generic vector $\omega$ in $\mathbb{C}^N$

is non-resonant, but for certain classes of vector fields like the class of Hamiltonian vector fields the spectrum vector is necessarily resonant—see Section 6.

As already mentioned, when $\omega$ is non-resonant, $F^\bullet$ is not necessarily trivial. This is because the alphabet $\mathcal{N} \subset \mathbb{C}^*$ is not necessarily stable under addition and it may happen that there is a nonempty word $\underline{\lambda} = \lambda_1 \cdots \lambda_r \in \underline{\mathcal{N}}$ such that $\lambda_1 + \cdots + \lambda_r = 0$, in which case formula (2.22) fails to define the value of $S^{\underline{\lambda}}$. In fact, in that case, there is no non-trivial mould $S^\bullet$ such that $\nabla S^\bullet = I^\bullet \times S^\bullet$. However, we repeat that *Poincaré's formal linearization theorem holds in that situation*: we necessarily have $B_{[\underline{\lambda}]} = 0$ for such a word $\underline{\lambda}$, and $Z = 0$, since there are no non-trivial resonant formal vector fields.

Here is an example in dimension $N = 2$: the spectrum vector $\omega = (5\varpi, 2\varpi)$ with $\varpi \in \mathbb{R}^*$ is non-resonant but if we assume that, associated with $(j,k) = \big(1,(0,2)\big)$ or $\big(1,(0,3)\big)$, there are nonzero coefficients $b_{j,k}$, then $\mathcal{N}$ contains $\lambda = -\varpi$ and $\mu = \varpi$ and (2.14)–(2.17) yield $F^{\lambda\mu} = \frac{1}{\varpi} = -F^{\mu\lambda}$ and $S^{\lambda\mu} = -\frac{1}{2\varpi^2} = S^{\mu\lambda}$.

**Remark 5.1.** If $\omega \in \mathbb{C}^N$ is "strongly non-resonant" in the sense that

$$\langle k, \omega\rangle \neq 0 \quad \text{for any nonzero } k \in \mathbb{Z}^N,$$

then the sum of the letters is nonzero for every nonempty word, hence $F^\bullet = 0$ and the symmetral mould $S^\bullet$ is entirely determined by the utterly simple formula (2.22). So, in *that* case, the mould equation $\nabla S^\bullet = I^\bullet \times S^\bullet$ has a symmetral solution, which is sufficient to obtain formal linearization by mould calculus.

**Remark 5.2.** On the other hand, it may happen that $\omega$ is resonant but 0 does not belong to the additive monoid generated by $\mathcal{N}$ (in particular this requires that $b_{j,k} = 0$ for every $(j,k) \in \mathcal{M}$ such that $\langle k,\omega\rangle - \omega_j = 0$). In that case $F^\bullet$ is necessarily 0, hence $X$ is formally linearizable.

**5.4** The formal flow $\Phi$ can be directly computed in terms of the symmetral mould $S^\bullet = \mathrm{e}^{G^\bullet}$: viewing the $B_\lambda$'s as differential operators which can be composed (and not only Lie-bracketed), we can define the associative comould $\underline{\lambda} = \lambda_1 \cdots \lambda_r \in \underline{\mathcal{N}} \mapsto B_{\lambda_1 \cdots \lambda_r} = B_{\lambda_r} \cdots B_{\lambda_1}$ and, according to the end of Remark 3.10, we get

$$Y = \sum_{r \geqslant 1} \sum_{\lambda_1,\dots,\lambda_r \in \mathcal{N}} G^{\lambda_1 \cdots \lambda_r} B_{\lambda_1 \cdots \lambda_r}$$

(in general $B_{\lambda_1 \cdots \lambda_r} \notin \mathcal{L}$, but the above sum is in $\mathcal{L}$ and coincides with $Y$), and

$$\mathrm{e}^Y = \mathrm{Id} + \sum_{r \geqslant 1} \sum_{\lambda_1,\dots,\lambda_r \in \mathcal{N}} S^{\lambda_1 \cdots \lambda_r} B_{\lambda_1 \cdots \lambda_r}$$

(this operator is not in $\mathcal{L}$). Now $\mathrm{e}^Y f = f\circ\Phi$ for any $f\in\mathbb{C}[[z_1,\dots,z_N]]$, hence $\Phi=(\Phi_1,\dots,\Phi_N)$ with

$$\Phi_j(z_1,\dots,z_N) = z_j + \sum_{r\geqslant 1}\ \sum_{\lambda_1,\dots,\lambda_r\in\mathcal{N}} S^{\lambda_1\cdots\lambda_r} B_{\lambda_1\cdots\lambda_r} z_j \quad \text{for } j=1,\dots,N.$$

There is a similar formula for $\Phi^{-1}$ involving ${}^{\mathrm{inv}}S^\bullet$.

## 6. **Classical Birkhoff normal forms**

**6.1** Let $d\in\mathbb{N}^*$. We now set

$$\mathcal{L}^{\mathbf{k}} := \big\{\, f\in\mathbf{k}[[x_1,\dots,x_d,y_1,\dots,y_d]] \mid f \text{ has order} \geqslant 2 \,\big\}, \qquad \mathbf{k}=\mathbb{R} \text{ or } \mathbb{C}.$$

The symplectic form $\sum_{j=1}^d \mathrm{d}x_j\wedge\mathrm{d}y_j$ induces the Poisson bracket $\{f,g\} := \sum_{j=1}^d \big(\frac{\partial f}{\partial x_j}\frac{\partial g}{\partial y_j} - \frac{\partial f}{\partial y_j}\frac{\partial g}{\partial x_j}\big)$, which makes $\mathcal{L}^{\mathbf{k}}$ a Poisson algebra over $\mathbf{k}$, and thus a Lie algebra over $\mathbf{k}$ with $[\cdot\,,\cdot]_{\mathrm{ham}} := \{\cdot\,,\cdot\}$. We get a complete filtered Lie algebra by setting $X\in\mathcal{L}^{\mathbf{k}}_{\geqslant m}$ if, as a power series, it has order $\geqslant m+2$.

Any $X\in\mathcal{L}^{\mathbf{k}}$ generates a formal Hamiltonian vector field, namely

$$\{X,\cdot\,\} = \sum_{j=1}^d \Big(\frac{\partial X}{\partial x_j}\frac{\partial}{\partial y_j} - \frac{\partial X}{\partial y_j}\frac{\partial}{\partial x_j}\Big)$$

viewed as a derivation of the associative algebra $\mathbf{k}[[x_1,\dots,x_d,y_1,\dots,y_d]]$. Let $X_0$ be the quadratic part of $X$, so that $\{X_0,\cdot\,\}$ is the linear part of the formal vector field $\{X,\cdot\,\}$. The corresponding matrix is Hamiltonian, hence its eigenvalues come into pairs of opposite complex numbers and we cannot avoid resonances in this case. From now on, we assume that

$$X_0 = \sum_{j=1}^d \tfrac12\omega_j(x_j^2+y_j^2), \quad \text{hence } \{X_0,\cdot\,\} = \sum_{j=1}^d \omega_j\Big(x_j\frac{\partial}{\partial y_j} - y_j\frac{\partial}{\partial x_j}\Big), \tag{6.1}$$

with a "frequency vector" $\omega=(\omega_1,\dots,\omega_d)\in\mathbf{k}^d$, so the eigenvalues of the linear part of the vector field are $\mathrm{i}\,\omega_1,\dots,\mathrm{i}\,\omega_d,-\mathrm{i}\,\omega_1,\dots,-\mathrm{i}\,\omega_d$ (which corresponds to a totally elliptic equilibrium point at the origin when $\mathbf{k}=\mathbb{R}$).

The formal Hamiltonian normalization problem consists in finding a formal symplectomorphism $\Phi$ such that the expression of $X\circ\Phi$ is as simple as possible (so that the expression of the conjugate Hamiltonian vector field $\Phi_*^{-1}\{X,\cdot\,\}$ is as simple as possible). We will apply Theorem A in the Lie algebra $\mathcal{L}^{\mathbb{C}}$ of complex formal Hamiltonian functions so as to recover the classical result according to which

> *there exists a formal symplectomorphism $\Phi$ (with real coefficients if $\mathbf{k}=\mathbb{R}$) such that $X\circ\Phi$ Poisson-commutes with $X_0$,*

i.e. $X \circ \Phi$ is a *Birkhoff normal form* (which implies, at the level of vector fields, that $\Phi_*^{-1}\{X, \cdot\}$ is a Hamiltonian Poincaré-Dulac normal form).

**6.2** The series

$$z_j(x,y) := \tfrac{1}{\sqrt{2}}(x_j + \mathrm{i}\, y_j), \quad w_j(x,y) := \tfrac{1}{\sqrt{2}}(\mathrm{i}\, x_j + y_j), \qquad j = 1, \dots, d, \tag{6.2}$$

satisfy $\sum \mathrm{d}x_j \wedge \mathrm{d}y_j = \sum \mathrm{d}z_j \wedge \mathrm{d}w_j$ and

$$\{X_0, z^k w^\ell\} = \mathrm{i}\,\langle k - \ell, \omega\rangle\, z^k w^\ell \quad \text{for any } k, \ell \in \mathbb{N}^d. \tag{6.3}$$

Using them as a change of coordinates and writing the generic formal series as

$$\sum_{k,\ell \in \mathbb{N}^d} b_{k,\ell}\, x^k y^\ell = \sum_{k,\ell \in \mathbb{N}^d} c_{k,\ell}\, z^k w^\ell,$$

we identify the complex Poisson algebras $\mathbb{C}[[x_1, \dots, x_d, y_1, \dots, y_d]]$ and $\mathbb{C}[[z_1, \dots, z_d, w_1, \dots, w_d]]$. The real Poisson algebra $\mathbb{R}[[x_1, \dots, x_d, y_1, \dots, y_d]]$ can be seen as the subspace consisting of the fixed points of the conjugate-linear involution $\mathscr{C}$ which maps $\sum b_{k,\ell}\, x^k y^\ell$ to $\sum \overline{b_{k,\ell}}\, x^\ell y^k$; note that $\mathscr{C}$ maps $\sum c_{k,\ell}\, z^k w^\ell$ to $\sum (-\mathrm{i})^{|k+\ell|} \overline{c_{k,\ell}}\, z^\ell w^k$, hence the coefficients $b_{k,\ell}$ are real if and only if

$$\overline{c_{k,\ell}} = \mathrm{i}^{|k+\ell|} c_{\ell,k} \quad \text{for all } k, \ell \in \mathbb{N}^d. \tag{6.4}$$

Let $X \in \mathcal{L}^{\mathbf{k}}$ with quadratic part $X_0$ as in (6.1). Introducing

$$\mathcal{M} := \big\{\, (k,\ell) \in \mathbb{N}^d \times \mathbb{N}^d \mid |k| + |\ell| \geqslant 3 \,\big\},$$

we can decompose $B := X - X_0 \in \mathcal{L}_1^{\mathbf{k}}$ as $B = \sum\limits_{(k,\ell) \in \mathcal{M}} c_{k,\ell}\, z^k w^\ell$ with coefficients $c_{k,\ell} \in \mathbb{C}$, and set

$$B_n := \sum_{\substack{(k,\ell) \in \mathcal{M} \text{ such} \\ \text{that } k - \ell = n}} c_{k,\ell}\, z^k w^\ell \in \mathcal{L}_1^{\mathbb{C}} \qquad \text{for } n \in \mathcal{N} := \mathbb{Z}^d, \tag{6.5}$$

so that $X = X_0 + \sum B_n$ and, for each $n \in \mathcal{N}$,

$$\{X_0, B_n\} = \lambda(n) B_n, \qquad \lambda(n) = \mathrm{i}\,\langle n, \omega\rangle \in \mathbb{C}. \tag{6.6}$$

Moreover, if $\mathbf{k} = \mathbb{R}$, then condition (6.4) holds, whence

$$\mathscr{C}(B_n) = B_{-n} \quad \text{for all } n \in \mathbb{Z}^d. \tag{6.7}$$

in that case.

**6.3** Let us apply Theorem A to $\mathcal{L}^{\mathbb{C}}$. For any complex-valued $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$ (recall that $\mathcal{N} = \mathbb{Z}^d$ and $\lambda$ is defined by (6.6)), Theorem B yields alternal moulds $F^\bullet, G^\bullet \in \mathbb{C}^{\underline{\mathcal{N}}}$, explicitly given by (2.14)–(2.17), such that $Z, Y \in \mathcal{L}^{\mathbb{C}}_{\geqslant 1}$ defined by

$$Z = \sum_{r \geqslant 1} \sum_{\underline{n} \in \mathcal{N}^r} \frac{1}{r} F^{\underline{n}} B_{[\underline{n}]}, \qquad Y = \sum_{r \geqslant 1} \sum_{\underline{n} \in \mathcal{N}^r} \frac{1}{r} G^{\underline{n}} B_{[\underline{n}]}$$

satisfy (1.2).

Formulas (2.14)–(2.17) show that, if $\mathbf{k} = \mathbb{R}$ and $A^\bullet$ is real-valued[6], then the complex conjugate of $F^{n_1\cdots n_r}$ is $F^{(-n_1)\cdots(-n_r)}$ and similarly for $G^\bullet$ (because $\overline{\lambda(n)} = \lambda(-n)$ for each $n \in \mathcal{N}$); on the other hand, $\mathscr{C}$ maps $B_{[\underline{n}]} = \{B_{n_r}, \{\ldots\{B_{n_2}, B_{n_1}\}\ldots\}\}$ to $\{B_{-n_r}, \{\ldots\{B_{-n_2}, B_{-n_1}\}\ldots\}\}$ (because of (6.7) and because $\mathscr{C}$ is a real Lie algebra automorphism[7] of $\mathcal{L}^{\mathbb{C}}$) and is conjugate-linear, hence we get $Z, Y \in \mathcal{L}^{\mathbb{R}}_{\geqslant 1}$ in that case.

So $Z, Y \in \mathcal{L}^{\mathbf{k}}_{\geqslant 1}$ whether $\mathbf{k} = \mathbb{C}$ or $\mathbb{R}$. The automorphism $\mathrm{e}^{\mathrm{ad}_Y}$ of $\mathcal{L}^{\mathbf{k}}$ is nothing but the action of the formal flow $\Phi$ at time 1 of the formal Hamiltonian vector field $\{Y, \cdot\}$ by composition: $\mathrm{e}^{\mathrm{ad}_Y} X = X \circ \Phi$, hence the second part of (1.2) says that $X \circ \Phi = X_0 + Z$, where $\Phi$ is a formal symplectomorphism with coefficients in $\mathbf{k}$, which implies $\Phi_*^{-1}\{X, \cdot\} = \{X_0 + Z, \cdot\}$ at the level of the formal Hamiltonian vector fields. The components of $\Phi$ can be directly computed from the symmetral mould $S^\bullet$ by means of (1.10):

$$\Phi_j(x,y) = x_j + \sum_{r\geqslant 1} \sum_{n_1,\ldots,n_r \in \mathcal{N}} S^{n_1\cdots n_r} \operatorname{ad}_{B_{n_r}} \cdots \operatorname{ad}_{B_{n_1}} x_j$$

$$\Phi_{d+j}(x,y) = y_j + \sum_{r\geqslant 1} \sum_{n_1,\ldots,n_r \in \mathcal{N}} S^{n_1\cdots n_r} \operatorname{ad}_{B_{n_r}} \cdots \operatorname{ad}_{B_{n_1}} y_j$$

for $j = 1, \ldots, N$ (the series $x_j$ and $y_j$ have been excluded from the definition of $\mathcal{L}^{\mathbf{k}}$, but (1.10) holds as an identity between operators acting in the whole of $\mathbf{k}[[x_1, \ldots, x_d, y_1, \ldots, y_d]]$).

The first part of (1.2) says that $X_0 + Z$ is a "Birkhoff normal form", in the sense that it Poisson-commutes with $X_0$. According to (6.3), this means that all the monomials in its $(z,w)$-expansion are of the form $c_{k,\ell}\, z^k w^\ell$ with $\langle k - \ell, \omega\rangle = 0$.

**6.4** Instead of (6.5), one can as well take

$$\mathcal{N} := \{\, \mathrm{i}\langle k-\ell, \omega\rangle \mid (k,\ell) \in \mathcal{M} \text{ and } c_{k,\ell} \neq 0 \,\} \subset \mathbb{C}, \qquad B_\lambda := \sum_{\substack{(k,\ell)\in\mathcal{M} \text{ such that}\\ \mathrm{i}\langle k-\ell,\omega\rangle = \lambda}} c_{k,\ell}\, z^k w^\ell,$$

so that (6.6) is replaced by $\{X_0, B_\lambda\} = \lambda B_\lambda$ for each $\lambda \in \mathcal{N}$ and one can use the formalism of Definition 2.4.

When $\omega$ is strongly non-resonant in the sense of Remark 5.1, the relation $\langle k - \ell, \omega\rangle = 0$ implies $k - \ell = 0$, hence

$$Z = \sum_{|\ell|\geqslant 2} C_\ell\, z^\ell w^\ell = \sum_{|\ell|\geqslant 2} \mathrm{i}^{|\ell|} C_\ell\, I_1^{\ell_1} \cdots I_d^{\ell_d}, \qquad I_j := \tfrac{1}{2}(x_j^2 + y_j^2) \quad \text{for } j = 1, \ldots, d,$$

[6] In fact it is sufficient that the complex conjugate of $A^{n_1\cdots n_r}$ is $A^{(-n_1)\cdots(-n_r)}$ for any word $n_1 \cdots n_r$.

[7] Indeed, $\mathscr{C}$ can be viewed as the symmetry $f_1 + \mathrm{i} f_2 \mapsto f_1 - \mathrm{i} f_2$ associated with the direct sum $\mathcal{L}^{\mathbb{C}} = \mathcal{L}^{\mathbb{R}} \oplus \mathrm{i}\mathcal{L}^{\mathbb{R}}$, it is a real Lie algebra automorphism because $\mathcal{L}^{\mathbb{R}}$ is a real Lie subalgebra.

with certain complex coefficients $C_\ell$, which satisfy $\mathrm{i}^{|\ell|}C_\ell \in \mathbb{R}$ when $\mathbf{k} = \mathbb{R}$.

It is easy to check that, when $\omega$ is strongly non-resonant, the Birkhoff normal form is unique (but not the formal symplectomorphism conjugating $X$ to it).

**6.5** *Remark.* Exactly the same formalism would apply to the perturbative situation of a Hamiltonian $X$ which is also a formal series in $\varepsilon$ (an indeterminate playing the role of a parameter). We would take $\mathcal{L}^{\mathbf{k}} := \mathbf{k}[[x_1, \dots, x_d, y_1, \dots, y_d, \varepsilon]]$ with $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$, with Lie bracket $[\cdot\,,\cdot]_{\mathrm{ham}} := \{\cdot\,,\cdot\}$ as before, and with filtration induced by the total order in the $2d+1$ indeterminates. Then, for any $X = X_0 + B$ with $X_0$ as in (6.1) and $B \in \mathcal{L}^{\mathbf{k}}_{\geqslant 1}$, Theorem A yields a formal symplectomorphism $\Phi$ such that $X \circ \Phi = X_0 + Z$ Poisson-commutes with $X_0$.

**6.6** The above formalism, as it stands, does not allow us to deal directly with $C^\infty$ functions of $(x, y)$, but there is a simple variant which allows for mixed Hamiltonians, formal in $\varepsilon$ (as in Remark 6.5) with coefficients $C^\infty$ in $(x, y)$. However, to have a decomposition of $X - X_0$ as a formally summable series of eigenvectors of $\{X_0, \cdot\}$, we must restrict ourselves to a certain kind of $C^\infty$ functions. With a view to allowing for comparison with certain quantum Hamiltonians in Section 9, we denote by $\mathcal{S}$ the Schwartz class and set, for $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$,

$$\mathscr{S}_0^{\mathbf{k}} := \Big\{ f \in \mathcal{S}(\mathbb{R}^d \times \mathbb{R}^d, \mathbf{k}) \mid \exists \tilde f \in C^\infty\big((\mathbb{R}_{\geqslant 0})^d, \mathbf{k}\big) \text{ such that } f(x,y) \equiv \tilde f\big(\tfrac{x_1^2+y_1^2}{2}, \dots, \tfrac{x_d^2+y_d^2}{2}\big) \Big\},$$

$$\mathscr{S}^{\mathbf{k}} := \Big\{ \sum_{(k,\ell)\in\Omega} b_{k,\ell}(x,y)\, x^k y^\ell \mid \Omega \text{ finite subset of } \mathbb{N}^d \times \mathbb{N}^d,\ b_{k,\ell} \in \mathscr{S}_0^{\mathbf{k}} \text{ for each } (k,\ell) \in \Omega \Big\},$$

$$\mathcal{L}^{\mathbf{k}} := \mathscr{S}^{\mathbf{k}}[[\varepsilon]].$$

We choose $\omega = (\omega_1, \dots, \omega_d) \in \mathbb{R}^d$ and consider the same $X_0$ as in (6.1). *Theorem A can be applied to any $X \in \mathcal{L}^{\mathbb{R}}$ of the form $X_0 + [\text{order} \geqslant 1 \text{ in } \varepsilon]$ so as to produce $Z, Y \in \mathcal{L}^{\mathbb{R}}$ such that $\{X_0, Z\} = 0$ and $\mathrm{e}^{\mathrm{ad}_Y} X = X_0 + Z$.*

Indeed, $\mathcal{L}^{\mathbb{R}}$ and $\mathcal{L}^{\mathbb{C}}$ are complete filtered Lie algebras (filtered by the order in $\varepsilon$), and $B := X - X_0$ can be decomposed into a formally convergent series as follows: we can write $B = \sum\limits_{\mathbb{N}^d\times\mathbb{N}^d} b_{k,\ell}(x,y,\varepsilon)\, x^k y^\ell$ with $b_{k,\ell}(x,y,\varepsilon) \in \mathscr{S}_0^{\mathbb{R}}[[\varepsilon]]_{\geqslant 1}$, hence $B = \sum_{n\in\mathbb{Z}^d} B_n$ with

$$B_n := \sum_{\substack{k',\ell',k'',\ell''\in\mathbb{N}^d \text{ such}\\ \text{that } k'+k''=n+\ell'+\ell''}} \frac{(-\mathrm{i})^{|\ell'+k''|}}{(\sqrt{2})^{|k'+k''+\ell'+\ell''|}} \binom{k'+\ell'}{k'} \binom{k''+\ell''}{k''} b_{k'+\ell',k''+\ell''}\, z(x,y)^{k'+k''} w(x,y)^{\ell'+\ell''}$$

with the same $z_j, w_j$ as in (6.2). This is the result of using $(x,y) \mapsto (z,w)$ as a change of coordinates; notice that the decomposition $B = \sum b_{k,\ell}(x,y,\varepsilon)\, x^k y^\ell$ is not unique, but the decomposition $B = \sum B_n$ is, and we have

$$\{X_0, B_n\} = \lambda(n) B_n, \qquad \lambda(n) = \mathrm{i}\langle n, \omega\rangle \in \mathbb{C} \qquad \text{for each } n \in \mathcal{N} := \mathbb{Z}^d.$$

Note that each $B_n \in \mathcal{L}^{\mathbb{C}}$, but the realness assumption on $X$ implies that $\mathscr{C}(B_n) = B_{-n}$ with the same conjugate-linear involution $\mathscr{C}$ as in Section 6.2. Therefore, for any real-valued $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, we get alternal moulds $F^\bullet, G^\bullet \in \mathbb{C}^{\underline{\mathcal{N}}}$ such that

$$Z = \sum_{r\geqslant 1} \sum_{\underline{n}\in\mathcal{N}^r} \frac{1}{r} F^{\underline{n}} B_{[\underline{n}]}, \qquad Y = \sum_{r\geqslant 1} \sum_{\underline{n}\in\mathcal{N}^r} \frac{1}{r} G^{\underline{n}} B_{[\underline{n}]}$$

define $Z, Y \in \mathcal{L}^{\mathbb{R}}_{\geqslant 1}$ with the desired properties (the realness of $Z$ and $Y$ follows from the same argument as in Section 6.3).

Note that if $\omega$ is strongly non-resonant in the sense of Remark 5.1, then $Z \in \mathscr{S}^{\mathbb{R}}_0[[\varepsilon]]$.

## 7. Multiphase averaging

**7.1** Let $d, N \in \mathbb{N}^*$. We call "slow-fast" a vector field of the form

$$X = \sum_{j=1}^{d} \big(\omega_j + \varepsilon f_j(\varphi, I, \varepsilon)\big) \frac{\partial}{\partial \varphi_j} + \sum_{k=1}^{N} \varepsilon g_k(\varphi, I, \varepsilon) \frac{\partial}{\partial I_k}, \tag{7.1}$$

where $\omega = (\omega_1, \ldots, \omega_d) \in \mathbb{R}^d$ is called the frequency vector, the idea being that, for $\varepsilon > 0$ "small", the time evolution of the variables $I_k$ will be "slow" compared to the "fast" variables $\varphi_j$ (at least if $\omega \neq 0$). We take $\varphi \in \mathbb{T}^d$, where $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, so the fast variables are angles. When $d = N$, this includes the case of vector field generated by a near-integrable Hamiltonian

$$X^{\mathrm{ham}} = \langle \omega, I \rangle + \varepsilon h(\varphi, I, \varepsilon) \tag{7.2}$$

for the symplectic form $\sum_{j=1}^d \mathrm{d}I_j \wedge \mathrm{d}\varphi_j$, for which $f_j = \frac{\partial h}{\partial I_j}$ and $g_j = -\frac{\partial h}{\partial \varphi_j}$.

We will deal with formal series in $\varepsilon$ whose coefficients are trigonometric polynomials in $\varphi$ with complex-valued coefficients smooth in $I$. More precisely, we take $f_1, \ldots, f_d, g_1, \ldots, g_N$ or $h$ in the complex associative algebra $\mathscr{A}^{\mathbb{C}}$ or the real associative algebra $\mathscr{A}^{\mathbb{R}}$ defined by

$$\mathscr{A}^{\mathbb{C}} := \mathscr{S}[\mathrm{e}^{\pm\mathrm{i}\,\varphi_1}, \ldots, \mathrm{e}^{\pm\mathrm{i}\,\varphi_d}][[\varepsilon]], \qquad \mathscr{A}^{\mathbb{R}} := \{\, f \in \mathscr{A}^{\mathbb{C}} \mid \overline{f(\overline{\varphi}, I, \overline{\varepsilon})} = f(\varphi, I, \varepsilon) \,\} \tag{7.3}$$

with $\mathscr{S} := C^\infty(D, \mathbb{C})$, where $D$ is an open subset of $\mathbb{R}^N$ (or $D = D' \times \mathbb{T}^{N''}$ with $D'$ open subset of $\mathbb{R}^{N'}$ and $N' + N'' = N$); in fact, we could as well take for $\mathscr{S}$ a linear subspace of $C^\infty(D, \mathbb{C})$, as long as it is stable under multiplication and all the derivations $\frac{\partial}{\partial I_k}$, e.g. one could take the Schwartz space $\mathcal{S}(\mathbb{R}^N, \mathbb{C})$.

Note that $\mathscr{A}^{\mathbb{R}}$ coincides with the set of fixed points of the conjugate-linear involution $\mathscr{C}$ which maps $\sum b_{n,p}(I)\, \varepsilon^p\, \mathrm{e}^{\mathrm{i}\langle n, \varphi\rangle}$ to $\sum \overline{b_{n,p}(I)}\, \varepsilon^p\, \mathrm{e}^{-\mathrm{i}\langle n, \varphi\rangle}$.

Let $X_0 := \sum \omega_j \frac{\partial}{\partial \varphi_j}$ and $X_0^{\mathrm{ham}} := \langle \omega, I \rangle$. *The formal averaging problem asks for a formal conjugacy between $X$ and a vector field $X_0 + Z$ which commutes with $X_0$ or, in the Hamiltonian* *version, for a formal symplectomorphism $\Phi$ such that $X^{\mathrm{ham}} \circ \Phi$ Poisson-commutes with $X_0^{\mathrm{ham}}$.* The reader is referred to [LM88] and [MS02] for the importance of this problem.

Let us set $\mathbf{k} := \mathbb{C}$ and consider the complete filtered Lie algebra $\mathcal{L}^{\mathbb{C}}$ consisting of vector fields whose components belong to $\mathscr{A}^{\mathbb{C}}$ (with $[\cdot\,,\cdot] = [\cdot\,,\cdot]_{\mathrm{vf}}$) or, in the Hamiltonian case, $\mathcal{L}^{\mathbb{C}} = \mathscr{A}^{\mathbb{C}}$ itself (with $[\cdot\,,\cdot] = \{\cdot\,,\cdot\}$, the Poisson bracket), filtered by the order in $\varepsilon$ in both cases. If we impose furthermore that the components of the vector fields or the Hamiltonian functions belong to $\mathscr{A}^{\mathbb{R}}$, then we get a real Lie subalgebra $\mathcal{L}^{\mathbb{R}}$.

**7.2** We can apply Theorem A to $\mathcal{L}^{\mathbb{C}}$. Indeed, any slow-fast system as above can be written as a sum of eigenvectors of $\mathrm{ad}_{X_0} = [X_0, \cdot\,]_{\mathrm{vf}}$ or $\mathrm{ad}_{X_0^{\mathrm{ham}}} = \{X_0^{\mathrm{ham}}, \cdot\,\}$,

$$X = X_0 + \sum_{n\in\mathcal{N}} B_n \quad \text{or} \quad X^{\mathrm{ham}} = X_0^{\mathrm{ham}} + \sum_{n\in\mathcal{N}} B_n^{\mathrm{ham}},$$

with $\mathcal{N} = \mathbb{Z}^d$ corresponding to all possible Fourier modes:

$$B_n = \mathrm{e}^{\mathrm{i}\langle n,\varphi\rangle}\Big(\sum_{j=1}^{d} b_{n,j}^{[1]}(I,\varepsilon)\frac{\partial}{\partial\varphi_j} + \sum_{k=1}^{N} b_{n,k}^{[2]}(I,\varepsilon)\frac{\partial}{\partial I_k}\Big), \qquad B_n^{\mathrm{ham}} = \mathrm{e}^{\mathrm{i}\langle n,\varphi\rangle} b_n(I,\varepsilon),$$

with certain coefficients $b_{n,j}^{[1]}, b_{n,k}^{[2]}, b_n \in \mathscr{S}[[\varepsilon]]$. In both cases, the eigenvalue map is

$$n \in \mathbb{Z}^d \mapsto \lambda(n) = \mathrm{i}\,\langle n,\omega\rangle \in \mathbb{C}. \tag{7.4}$$

For any choice of $A^\bullet \in \mathrm{Alt}^\bullet_{\lambda=0}(\mathcal{N})$, we thus get $Y, Z \in \mathcal{L}^{\mathbb{C}}$ of order $\geqslant 1$ in $\varepsilon$ such that

$$[X_0, Z]_{\mathrm{vf}} = 0 \;\text{ and }\; \mathrm{e}^{\mathrm{ad}_Y} X = X_0 + Z, \quad \text{or} \quad \{X_0^{\mathrm{ham}}, Z\} = 0 \;\text{ and }\; \mathrm{e}^{\mathrm{ad}_Y} X^{\mathrm{ham}} = X_0^{\mathrm{ham}} + Z.$$

In the first case, as in Section 5,

$$\mathrm{e}^{\mathrm{ad}_Y} X = \Phi_*^{-1} X \tag{7.5}$$

where $\Phi$ is the formal flow at time 1 of the formal vector field $Y$. In the second case, as in Section 6,

$$\mathrm{e}^{\mathrm{ad}_Y} X^{\mathrm{ham}} = X^{\mathrm{ham}} \circ \Phi \tag{7.6}$$

where $\Phi$ is the formal symplectomorphism obtained by flowing at time 1 along the formal Hamiltonian vector field $\{Y, \cdot\,\}$. In both cases,

$$Z \text{ only contains Fourier modes } n \in \mathcal{N} \text{ such that } \langle n,\omega\rangle = 0. \tag{7.7}$$

Therefore, when $\omega$ is strongly non-resonant in the sense of Remark 5.1, the components of the formal vector field $Z$ (in the first case) or the formal series $Z$ (in the second case) do not depend on $\varphi$, they are formal series in $\varepsilon$ with coefficients depending on $I$ only: *the formal change of coordinates $\Phi^{-1}$ has eliminated the fast phase $\varphi$ from the vector field.*

If the coefficients $f_1, \dots, f_d, g_1, \dots, g_N$ or $h$ belong to $\mathscr{A}^{\mathbb{R}}$, i.e. if we start with $X$ or $X^{\mathrm{ham}}$ in $\mathcal{L}^{\mathbb{R}}$, and we take $A^\bullet$ real-valued, then one gets $Y, Z \in \mathcal{L}^{\mathbb{R}}$ for the same reason as in Section 6: $\mathcal{L}^{\mathbb{R}}$ consists of the fixed points of $\mathscr{C}$ which is a real Lie algebra automorphism[8] mapping $B_n$ to $B_{-n}$ and is conjugate-linear, and the complex conjugate of $F^{n_1 \cdots n_r}$ is $F^{(-n_1)\cdots(-n_r)}$ and similarly for $G^\bullet$ (the condition described in footnote 6 is sufficient for this).

**7.3** *Remark.* In the real Hamiltonian case, $X_0^{\mathrm{ham}} + Z$ can be considered as a Birkhoff normal form for $X^{\mathrm{ham}} = X_0^{\mathrm{ham}} + \varepsilon h(\varphi, I, \varepsilon)$. If we choose $\mathscr{S} = \mathcal{S}(\mathbb{R}^N, \mathbb{C})$ in (7.3), then we get the action-angle analogue of Section 6.6.

## 8. **Quantum Birkhoff normal forms**

**8.1** Let $\mathcal{H}$ be a complex Hilbert space, with inner product denoted by $\langle \cdot \,|\, \cdot \rangle$. In this section, by "operator", we mean an unbounded linear operator with dense domain.

Let us consider an operator $X_0$ of $\mathcal{H}$ which is diagonal in an orthonormal basis $\mathbf{e} = (e_k)_{k \in I}$ of $\mathcal{H}$:

$$X_0\, e_k = E_k\, e_k, \qquad k \in I,$$

with eigenvalues $E_k \in \mathbb{C}$, i.e. $X_0$ is a normal operator, or $E_k \in \mathbb{R}$, in which case $X_0$ is self-adjoint. Let $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$ consist of all operators of $\mathcal{H}$ whose domain is the dense subspace $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$ and which preserve $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$. Let $\mathcal{L}_{\mathbf{e}}^{\mathbb{R}}$ consist of all symmetric operators among the previous ones. In particular, the restriction of $X_0$ to $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$ belongs to $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$, and even to $\mathcal{L}_{\mathbf{e}}^{\mathbb{R}}$ in the self-adjoint case.

Notice that an element $B$ of $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$ is determined by a complex "infinite matrix" $(\beta_{k,\ell})_{k,\ell \in I}$:

$$B e_k = \sum_{\ell \in I} \beta_{k,\ell}\, e_\ell, \qquad k \in I, \tag{8.1}$$

with the following "finite-column" property: if $\beta_{k,\ell} \neq 0$ then $\ell$ belongs to a finite subset of $I$ depending on $k$ and $B$. The domain of the adjoint operator $B^*$ then contains $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$, and

$$B^* e_k = \sum_{\ell \in I} \overline{\beta_{\ell,k}}\, e_\ell, \qquad k \in I.$$

[8] To see it, first observe that $C\colon (\varphi, I) \mapsto (-\varphi, I)$ is conformal-symplectic with a factor $-1$ hence the composition with $C$ is a complex Lie algebra anti-automorphism $\Theta_C$ of $\mathcal{L}^{\mathbb{C}}$, then note that $\mathscr{C} = \Theta_C \circ \mathbf{S}$ where $\mathbf{S}$ is the symmetry associated with the direct sum $\mathscr{A}^{\mathbb{C}} = \mathscr{R} \oplus \mathrm{i}\mathscr{R}$, with $\mathscr{R} := C^\infty(D, \mathbb{R})[\mathrm{e}^{\pm \mathrm{i}\varphi_1}, \dots, \mathrm{e}^{\pm \mathrm{i}\varphi_d}][[\varepsilon]]$ real linear subspace, and $\mathbf{S}$ is a real Lie algebra anti-automorphism because the Lie bracket of vector fields with components in $\mathscr{R}$ has its components in $\mathrm{i}\mathscr{R}$ and, for Hamiltonians, $\{\mathscr{R}, \mathscr{R}\} \subset \mathrm{i}\mathscr{R}$.

**Lemma 8.1.** *(i) For $A, B \in \mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$, there is a well-defined composite operator $AB \in \mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$, and for this product $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$ is an associative algebra over $\mathbb{C}$.*
*(ii) Let $\hbar > 0$ be fixed. The formula*

$$[A, B]_{\mathrm{qu}} := \tfrac{1}{\mathrm{i}\hbar}(AB - BA), \qquad A, B \in \mathscr{A}_{\mathbf{e}}^{\mathbb{C}},$$

*makes $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$ a Lie algebra over $\mathbb{C}$, which we denote by $\mathcal{L}_{\mathbf{e}}^{\mathbb{C}}$.*
*(iii) $\mathcal{L}_{\mathbf{e}}^{\mathbb{R}}$ is a real Lie subalgebra of $\mathcal{L}_{\mathbf{e}}^{\mathbb{C}}$, coinciding with the set of the fixed points of the involution*

$$\mathscr{C} : \ B \in \mathscr{A}_{\mathbf{e}}^{\mathbb{C}} \mapsto B^*{}_{|\,\mathrm{Span}_{\mathbb{C}}(\mathbf{e})} \in \mathscr{A}_{\mathbf{e}}^{\mathbb{C}},$$

*which is a conjugate-linear anti-homomorphism of the associative algebra $\mathscr{A}_{\mathbf{e}}^{\mathbb{C}}$, and a real Lie algebra automorphism of $\mathcal{L}_{\mathbf{e}}^{\mathbb{C}}$.*

*Proof.* Obvious. □

**8.2** We want to perturb $X_0$ in $\mathcal{L}_{\mathbf{e}}^{\mathbb{C}}$, resp. in $\mathcal{L}_{\mathbf{e}}^{\mathbb{R}}$, by a "small" perturbation and work formally, as in a Rayleigh-Schrödinger-like situation. So, we introduce an indeterminate $\varepsilon$ and consider

$$\mathcal{L}^{\mathbb{C}} := \mathcal{L}_{\mathbf{e}}^{\mathbb{C}}[[\varepsilon]], \quad \text{resp.} \quad \mathcal{L}^{\mathbb{R}} := \mathcal{L}_{\mathbf{e}}^{\mathbb{R}}[[\varepsilon]],$$

as a complete filtered Lie algebra over $\mathbb{C}$, resp. over $\mathbb{R}$, filtered by order in $\varepsilon$.

To decompose an arbitrary perturbation as a sum of eigenvectors of $\mathrm{ad}_{X_0}$, we notice that, for $B \in \mathcal{L}^{\mathbb{C}}$ with matrix $\big(\beta_{k,\ell}(\varepsilon)\big)_{k,\ell \in I}$ so that (8.1) holds (with formal series $\beta_{k,\ell}(\varepsilon) \in \mathbb{C}[[\varepsilon]]$), we can write

$$B = \sum_{(k,\ell) \in I \times I} \tilde{B}_{k,\ell} \quad \text{with } \tilde{B}_{k,\ell} := |\, e_\ell \rangle \, \beta_{k,\ell}(\varepsilon) \, \langle e_k \,| \tag{8.2}$$

(here we used the Dirac notation i.e. $\tilde{B}_{k,\ell} e_j = \beta_{k,\ell}(\varepsilon)\, e_\ell$ if $j = k$, $\tilde{B}_{k,\ell} e_j = 0$ else). The sum in (8.2) may be infinite, but it is well-defined because its action in $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$ is finitary. One then easily checks that

$$\Big[X_0, \tilde{B}_{k,\ell}\Big]_{\mathrm{qu}} = \tfrac{1}{\mathrm{i}\hbar}(E_\ell - E_k)\tilde{B}_{k,\ell}.$$

We thus have $B = \sum_{\lambda \in \mathcal{N}} B_n$ with

$$\mathcal{N} := \big\{ \tfrac{1}{\mathrm{i}\hbar}(E_\ell - E_k) \mid (k,\ell) \in I \times I \big\}, \qquad B_\lambda := \sum_{\substack{(k,\ell) \text{ such that} \\ E_\ell - E_k = \mathrm{i}\hbar\lambda}} |\, e_\ell \rangle \beta_{k,\ell}(\varepsilon) \langle e_k \,| \quad \text{for } \lambda \in \mathcal{N}. \tag{8.3}$$

Note that, if $X_0, B \in \mathcal{L}^{\mathbb{R}}$, then

$$\mathscr{C}(B_\lambda) = B_{-\lambda} \quad \text{for any } \lambda \in \mathcal{N}. \tag{8.4}$$

We thus suppose that we are given a perturbation $B \in \mathcal{L}^{\mathbb{C}}_{\geqslant 1}$. We can apply Theorem A to $X = X_0 + B \in \mathcal{L}^{\mathbb{C}}$, with $\mathbf{k} = \mathbb{C}$. For each choice of $A^\bullet \in \mathrm{Alt}^\bullet_0(\mathcal{N})$, we get $Z, Y \in \mathcal{L}^{\mathbb{C}}$ of order $\geqslant 1$ in $\varepsilon$ such that

$$[X_0, Z]_{\mathrm{qu}} = 0, \qquad \mathrm{e}^{\mathrm{ad}_Y} X = X_0 + Z. \tag{8.5}$$

Since $\mathscr{A}^{\mathbb{C}}_{\mathbf{e}}[[\varepsilon]]$ is a complete filtered associative algebra and $Y$ is of order $\geqslant 1$ in $\varepsilon$, we can define $U := \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Y}$ by the exponential series: it is an automorphism of $\mathrm{Span}_{\mathbb{C}}(\mathbf{e})$ formal in $\varepsilon$, with inverse $U^{-1} = \mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}Y}$, and $\mathrm{e}^{\mathrm{ad}_Y} X = U X U^{-1}$. So, the second part of (8.5) says that

$$U(X_0 + B)U^{-1} = X_0 + Z, \qquad U = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Y}.$$

Mould calculus shows that

$$\tfrac{1}{\mathrm{i}\hbar} Y = \sum_{r \geqslant 1} \sum_{\lambda_1, \ldots, \lambda_r \in \mathcal{N}} (\tfrac{1}{\mathrm{i}\hbar})^r G^{\lambda_1 \cdots \lambda_r} B_{\lambda_r} \cdots B_{\lambda_1}, \qquad U = \mathrm{Id} + \sum_{r \geqslant 1} \sum_{\lambda_1, \ldots, \lambda_r \in \mathcal{N}} (\tfrac{1}{\mathrm{i}\hbar})^r S^{\lambda_1 \cdots \lambda_r} B_{\lambda_r} \cdots B_{\lambda_1},$$

and there is a similar formula for $U^{-1}$ involving the mould ${}^{\mathrm{inv}}S^\bullet$.

If we assume that each eigenvalue $E_k$ of $X_0$ is simple (an assumption analogous to the strong non-resonance condition of Remark 5.1), then it is easy to check that the first part of (8.5) says that $Z$ is diagonal in the basis $\mathbf{e}$. In general, it says that $Z$ is block-diagonal, where the blocks refer to the partition $I = \bigsqcup I_a$, $I_a := \{\, k \in I \mid E_k = a \,\}$.

Suppose now that $X_0 \in \mathcal{L}^{\mathbb{R}}$, i.e. it is a self-adjoint operator, and also $B \in \mathcal{L}^{\mathbb{R}}$. Then, in view of (8.4), by the same arguments as in Section 6 or 7, we get $Z, Y \in \mathcal{L}^{\mathbb{R}}$. Note that $U$ is then a "formal unitary operator". The formally conjugate operator $X_0 + Z$ is called a quantum Birkhoff normal form for $X_0 + B$.

**8.3** The simplest example is that of the self-adjoint operator $X_0 = -\mathrm{i}\hbar \sum \omega_j \frac{\partial}{\partial \varphi_j}$ of $\mathcal{H} = L^2(\mathbb{T}^d)$, which is diagonal in the Fourier basis. We have $I = \mathbb{Z}^d$ and, for each $k \in \mathbb{Z}^d$, $e_k = (2\pi)^{-d/2}\, \mathrm{e}^{\mathrm{i}\langle k, \varphi \rangle}$ and the corresponding eigenvalue is $E_k = \hbar\, \langle k, \omega \rangle$ for $k \in \mathbb{Z}^d$. In particular,

$$\tfrac{1}{\mathrm{i}\hbar}(E_\ell - E_k) = \mathrm{i}\, \langle k - \ell, \omega \rangle.$$

The simplest example for $\mathcal{H} = L^2(\mathbb{R}^d)$ is the quantum harmonic oscillator

$$X_0 = -\tfrac{1}{2}\hbar^2 \Delta + \sum_{j=1}^{d} \tfrac{1}{2} \omega_j^2 x_j^2 \tag{8.6}$$

(with $\omega_1, \ldots, \omega_d > 0$ given), for which the spectrum is natually indexed by $I = \mathbb{N}^d$:

$$E_k = \hbar\, \langle k + (\tfrac{1}{2}, \ldots, \tfrac{1}{2}), \omega \rangle, \qquad k \in \mathbb{N}^d, \tag{8.7}$$

and $\mathbf{e}$ is given by the Hermite functions.

In these cases, one can index the eigenvector decomposition $B = \sum B_n$ of finite-column operators by $\mathcal{N} = \mathbb{Z}^d$, by a slight modification of (8.3):

$$B_n := \sum_{\substack{(k,\ell)\in\mathbb{N}^d\times\mathbb{N}^d \\ k-\ell=n}} |\, e_\ell \rangle\, \beta_{k,\ell}(\varepsilon)\, \langle e_k \,|, \qquad n \in \mathbb{Z}^d.$$

This way, the eigenvalue map is $\lambda(n) = \mathrm{i}\,\langle n, \omega\rangle$.

Moreover, in these cases, one may wish to restrict oneself to the "finite-band" case defined by replacing $\mathcal{L}^{\mathbb{R}}_{\mathbf{e}}$ with its subspace $\mathcal{L}^{\mathbb{R}}_{\mathbf{e},\mathrm{fb}}$ consisting of those elements associated with infinite matrices $(\beta_{k,\ell})_{k,\ell\in I}$ for which there exists $K \in \mathbb{N}$ such that $\beta_{k,\ell} = 0$ for $|k-\ell| < K$. Since $\mathcal{L}^{\mathbb{R}}_{\mathrm{fb}} := \mathcal{L}^{\mathbb{R}}_{\mathbf{e},\mathrm{fb}}[[\varepsilon]]$ is a Lie subalgebra of $\mathcal{L}^{\mathbb{R}}$, we get $Z, Y \in \mathcal{L}^{\mathbb{R}}_{\mathrm{fb}}$ whenever we start with a perturbation $B \in \mathcal{L}^{\mathbb{R}}_{\mathrm{fb}}$ or order $\geqslant 1$ in $\varepsilon$.

## 9. **Semi-classical limit**

**9.1** In general the dependence of the eigenvalues $E_k$ in the Planck constant $\hbar$ is very complicated, very often intractable. This makes the set $\mathcal{N} = \mathcal{N}(\hbar)$ in (8.3) very difficult to follow as $\hbar \to 0$. Nevertheless, this difficulty is absent in the two examples of $X_0$ of Section 8.3, since we have seen that in these cases we can choose $\mathcal{N} = \mathbb{Z}^d$ and $\lambda(n) = \mathrm{i}\,\langle n, \omega\rangle$, thus independent of $\hbar$.

We will now consider an operator $X = X_0 + B^{\mathrm{qu}}$ obtained by Weyl quantization[9] from a classical Hamiltonian $\sigma(x,\xi,\varepsilon)$ of the type introduced in Section 6.6. For the sake of simplicity, we choose $X_0$ to be the quantum harmonic oscillator (8.6) on $L^2(\mathbb{R}^d)$ (we could treat as well the case of the trickier Weyl quantization on $\mathbb{T}^d$ and choose for $X_0$ the first example of Section 8.3, starting from a classical Hamiltonian $\sigma(x,\xi,\varepsilon)$ of the type alluded to in Section 7.3). We take arbitrary $\omega_1,\dots,\omega_d > 0$; it will *not* be necessary to assume that the corresponding frequency vector $\omega := (\omega_1,\dots,\omega_d)$ is non-resonant.

The quantum harmonic oscillator $X_0$ is the Weyl quantization of the Hamiltonian

$$\sigma_0(x,\xi) := \sum_{j=1}^{d} \tfrac{1}{2}(\xi_j^2 + \omega_j^2 x_j^2), \tag{9.1}$$

which differs from the quadratic Hamiltonian (6.1) considered in Section 6 only by the conformal-symplectic change of coordinates induced by $\xi_j = \omega_j y_j$. Let us thus consider a formal Hamiltonian $\sigma \in \mathscr{S}^{\mathbb{R}}[[\varepsilon]]$ of the form

$$\sigma = \sigma_0 + B^{\mathrm{cl}}, \quad \text{with } B^{\mathrm{cl}} = B^{\mathrm{cl}}(x,\xi,\varepsilon) \text{ of order} \geqslant 1 \text{ in } \varepsilon, \tag{9.2}$$

[9]See e.g. [Fol89] for a general exposition of pseudo-differential operators and Weyl quantization. The few definitions and facts we need will be recalled in Section 9.2.

exactly as in Section 6.6 except for the change $y \to \xi$. Weyl quantization gives rise to a self-adjoint operator $X = X_0 + B^{\rm qu}$ of $L^2(\mathbb{R}^d)$. We are interested in comparing the quantum Birkhoff normal form $X_0 + Z^{\rm qu}$ of $X$ and the classical Birkhoff normal form $\sigma_0 + Z^{\rm cl}$ of $\sigma$.

We will see how transparent mould calculus makes the relation between $Z^{\rm qu}$ and $Z^{\rm cl}$. The point is that it is the very same mould $F^\bullet$ which will appear in the mould expansions $Z^{\rm cl} = F^\bullet B^{\rm cl}_{[\bullet]}$ and $Z^{\rm qu} = F^\bullet B^{\rm qu}_{[\bullet]}$; the difference lies only in the Lie comould to be used in each expansion, but the semi-classical limit of the quantum Lie comould $B^{\rm qu}_{[\bullet]}$ is easily tractable in this context, with its symbol tending to $B^{\rm cl}_{[\bullet]}$ as $\hbar \to 0$. In fact, all the "difficult" part, that is solving the mould equation which generates combinatorial difficulties solved only by induction, is exactly the same in the classical and quantum cases.

**9.2** The operator $X_0$ is obtained from $\sigma_0$ by replacing $\xi_j$ by $-i\hbar\frac{\partial}{\partial x_j}$. More generally, Weyl quantization associates to a function $\sigma$ belonging e.g. to the Schwartz class $\mathcal{S}(\mathbb{R}^d \times \mathbb{R}^d) = \mathcal{S}(T^*\mathbb{R}^d)$ an operator $\mathcal{V}$ which acts on a function $\varphi \in L^2(\mathbb{R}^d)$ through the formula

$$\mathcal{V}\varphi(x) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \sigma\Big(\frac{x+y}{2}, \xi\Big)\, \mathrm{e}^{\mathrm{i}\frac{\xi(x-y)}{\hbar}}\, \varphi(y)\frac{\mathrm{d}\xi\mathrm{d}y}{(2\pi\hbar)^d}. \tag{9.3}$$

In other words, the operator $\mathcal{V}$ has an integral kernel given by

$$K_{\mathcal{V}}(x,y) := \int_{\mathbb{R}^d\times\mathbb{R}^d} \sigma\Big(\frac{x+y}{2}, \xi\Big)\, \mathrm{e}^{\mathrm{i}\frac{\xi(x-y)}{\hbar}}\, \frac{\mathrm{d}\xi}{(2\pi\hbar)^d}.$$

A straightforward computation shows that this formula is invertible by

$$\sigma(x,\xi) = \int_{\mathbb{R}^d} K_{\mathcal{V}}(x+\delta, x-\delta)\, \mathrm{e}^{-2\mathrm{i}\frac{\xi\delta}{\hbar}}\, \mathrm{d}\delta. \tag{9.4}$$

In that situation, we use the notation $\sigma = \sigma_{\mathcal{V}}$ and say that the function $\sigma$ is the "symbol" of the operator $\mathcal{V}$. For instance, with the notations of Section 9.1, $\sigma_0 = \sigma_{X_0}$.

The following result is the fundamental one concerning the transition quantum-classical. Its proof is straightforward for symbols in the Schwartz class, by using (9.3) and (9.4). It gives a mod($\hbar$)-homomorphism between quantum and classical Lie algebras.

**Lemma 9.1.** *Suppose that the operators $V$ and $W$ are obtained by Weyl quantization from the symbols $\sigma_V$ and $\sigma_W$. Then the symbol of $\frac{1}{\mathrm{i}\hbar}[W,V]$ is*

$$\sigma_{\frac{1}{\mathrm{i}\hbar}[W,V]} = A(\sigma_W \otimes \sigma_V), \tag{9.5}$$

*where* $A(f\otimes g)(x,\xi) = \frac{1}{\hbar}\sin\Big(\hbar\big(\frac{\partial}{\partial q}\frac{\partial}{\partial p'} - \frac{\partial}{\partial p}\frac{\partial}{\partial q'}\big)\Big) f(q,p)g(q',p')|_{q=q'=x,\, p=p'=\xi}$.

*In particular*

$$\lim_{\hbar\to 0} \sigma_{\frac{1}{\mathrm{i}\hbar}[W,V]} = \{\sigma_W, \sigma_V\} \tag{9.6}$$

*and, in the case of a quadratic symbol* $\sigma_{X_0}$ *like in* (9.1)*,*

$$\sigma_{\frac{1}{\mathrm{i}\hbar}[X_0,V]} = \{\sigma_{X_0}, \sigma_V\}. \tag{9.7}$$

**9.3** On the one hand, according to Section 6.6, the Hamiltonian (9.2) can be decomposed as

$$\sigma = \sigma_0 + \sum_{n\in\mathcal{N}} B_n^{\mathrm{cl}}, \qquad \{\sigma_0, B_n^{\mathrm{cl}}\} = \lambda(n) B_n^{\mathrm{cl}}, \qquad \lambda(n) = \mathrm{i}\langle n, \omega\rangle,$$

with $\mathcal{N} := \mathbb{Z}^d$. Denoting by $B_{[\bullet]}^{\mathrm{cl}}$ the Lie comould defined from $(B_n^{\mathrm{cl}})_{n\in\mathcal{N}}$ by means of Poisson brackets, we get a Birkhoff normal form of $\sigma$ in the form $\sigma_0 + Z^{\mathrm{cl}}$ with

$$Z^{\mathrm{cl}} = \sum_{r\geqslant 1} \sum_{\underline{n}\in\mathcal{N}^r} \frac{1}{r} F^{\lambda(n_1)\cdots\lambda(n_r)}\, B_{[\underline{n}]}^{\mathrm{cl}}, \tag{9.8}$$

where we choose for $F^\bullet$ the first of a pair of alternal moulds $(F^\bullet, G^\bullet)$ solving (2.10) in the canonical case of Section 2.5.2 (we may choose any alternal solution, e.g. the zero gauge solution; note that if $\omega$ is strongly non-resonant, then $Z^{\mathrm{cl}}$ is uniquely determined, hence this choice is not relevant, but we make no such hypothesis about $\omega$).

On the other hand, the Weyl quantization of $\sigma = \sigma_0 + B^{\mathrm{cl}}$ is $X = X_0 + B^{\mathrm{qu}}$ and, for each $n \in \mathcal{N}$, the Weyl quantization $B_n^{\mathrm{qu}}$ of $B_n^{\mathrm{cl}}$ satisfies

$$\sigma_{\frac{1}{\mathrm{i}\hbar}[X_0,B_n^{\mathrm{qu}}]} = \{\sigma_0, B_n^{\mathrm{cl}}\} = \sigma_{\lambda(n)B_n^{\mathrm{qu}}}$$

because of (9.7), hence $B_n^{\mathrm{qu}}$ is the $n$-homogeneous component of $B^{\mathrm{qu}}$. Note that $B^{\mathrm{qu}}$ and the $B_n^{\mathrm{qu}}$'s belong to the space $\mathcal{L}_{\mathbf{e},\mathrm{fb}}^{\mathbb{R}}[[\varepsilon]]$ defined at the end of Section 8.3. Now, according to Section 8, we obtain a quantum Birkhoff normal form of $X$ in the form $X_0 + Z^{\mathrm{qu}}$ with

$$Z^{\mathrm{qu}} = \sum_{r\geqslant 1} \sum_{\underline{n}\in\mathcal{N}^r} \frac{1}{r} F^{\lambda(n_1)\cdots\lambda(n_r)}\, B_{[\underline{n}]}^{\mathrm{qu}}, \tag{9.9}$$

if we take for $F^\bullet$ the *same* mould as in (9.8) and define $B_{[\bullet]}^{\mathrm{qu}}$ as the Lie comould generated by $(B_n^{\mathrm{qu}})_{n\in\mathcal{N}}$ by means of the Lie bracket $[\cdot\,,\cdot]_{\mathrm{qu}}$ of $\mathcal{L}_{\mathbf{e},\mathrm{fb}}^{\mathbb{R}}[[\varepsilon]]$ (note that, *if* $\omega$ is strongly non-resonant, then the eigenvalues (8.7) are simple and $Z^{\mathrm{qu}}$ is uniquely determined).

For each letter $n \in \mathcal{N}$, the symbol of $B_n^{\mathrm{qu}}$ is the Hamiltonian $B_n^{\mathrm{cl}}$, but in general, for a word $\underline{n} \in \underline{\mathcal{N}}$ of length $\geqslant 2$, the symbol of $B_{[\underline{n}]}^{\mathrm{qu}}$ is not exactly $B_{[\underline{n}]}^{\mathrm{cl}}$. However, iteration of (9.6) implies

$$\lim_{\hbar\to 0} \sigma_{B_{[\underline{n}]}^{\mathrm{qu}}} = B_{[\underline{n}]}^{\mathrm{cl}} \quad \text{for each nonempty } \underline{n} \in \underline{\mathcal{N}}. \tag{9.10}$$

Putting together (9.8), (9.9) and (9.10), we thus obtain very simply the following result:

**Theorem.** *One has*

$$\sigma_{Z^{\mathrm{qu}}} \xrightarrow[\hbar\to 0]{} Z^{\mathrm{cl}} \quad \textit{termwise in } \varepsilon,$$

*i.e. the coefficients of the $\varepsilon$-expansion of the classical Birkhoff normal form $X_0 + Z^{\rm cl}$ are the limits, as $\hbar \to 0$, of the symbols of the coefficients of the $\varepsilon$-expansion of the quantum Birkhoff normal form $X_0 + Z^{\rm qu}$.*

In the case of a strongly non-resonant frequency vector $\omega$ satisfying a Diophantine condition, this result was first established in [GP87] and later using the Lie method in [DGH91].

**Acknowledgments**: The authors are grateful to Fréderic Menous for pointing out the reference [Men13]. This work has been partially carried out thanks to the support of the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR). The research leading these results was also partially supported by the French National Research Agency under the reference ANR-12-BS01-0017. T.P. thanks the Dipartimento di Matematica, Sapienza Università di Roma, for its kind hospitality during the completion of this work.

## References

[DGH91] M. Degli Esposti, S. Graffi, J. Herczynski, *Quantization of the classical Lie algorithm in the Bargmann representation*, Annals of Physics, **209** 2, (1991) 364–392.

[Eca81] J. Écalle, *Les fonctions résurgentes*, Publ. Math. d'Orsay [Vol. 1: 81-05, Vol. 2: 81-06, Vol. 3: 85-05] 1981, 1985.

[Eca93] J. Écalle, Six lectures on Transseries, Analysable Functions and the Constructive Proof of Dulac's conjecture, in *Bifurcations and periodic orbits of vector fields (Montreal, PQ, 1992)* (ed. by D. Schlomiuk), NATO Adv. Sci. Inst. Ser.C Math. Phys. Sci. 408, Kluwer Acad. Publ., Dordrecht 1993, 75–184.

[EV95] J. Écalle and B. Vallet, *Prenormalization, correction, and linearization of resonant vector fields or diffeomorphisms.* Prepub. Orsay 95-32 (1995), 90 pp.

[Fol89] G. Folland, *Harmonic Analysis in Phase Space*, Annals of Mathematics Studies **122**, Princeton University Press 1989.

[GP87] S. Graffi, T.Paul, *Schrödinger equation and canonical perturbation theory*, Comm. Math. Phys., **108** (1987), 25–40.

[LM88] P. Lochak, C. Meunier, *Multiphase averaging for classical systems*, Applied Mathematical Sciences, **72**, Springer-Verlag, New York, 1988, xii+360 pp.

[MS02] J.-P. Marco, D. Sauzin, "Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* **96** (2002), 199–275.

[Men13] F. Menous, *From dynamical systems to renormalization*, Journal of Mathematical Physics **54** (2013), 092702 1-24.

[PS16] T. Paul and D. Sauzin, *Normalization in Banach scales of Lie algebras via mould calculus and applications*, in preparation.

[Sau09] D. Sauzin, "Mould expansions for the saddle-node and resurgence monomials," in *Renormalization and Galois theories*, p. 83–163, A. Connes, F. Fauvet, J.-P. Ramis (eds.), IRMA Lectures in Mathematics and Theoretical Physics 15, Zürich: European Mathematical Society, 2009.



CMLS, Ecole polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau Cedex, France
*E-mail address*: thierry.paul@polytechnique.edu

CNRS UMI 8028–IMCCE, 77 av. Denfert-Rochereau, 75014 Paris, France,
CNRS UMI 3483–Laboratorio Fibonacci, CRM E. De Giorgi, Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 3, 56126 Pisa, Italy
*E-mail address*: David.Sauzin@obspm.fr